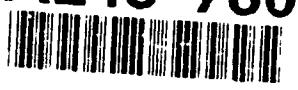


AN INVESTIGATION OF THE APPLICATION OF ARTIFICIAL NEURAL NETWORKS TO ADAPTIVE OPTICS IMAGING SYSTEMS

THESIS

Andrew H. Suzuki
Captain, USAF

AFIT/GEO/ENG/91D-04

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1991	*Master's Thesis*

4. TITLE AND SUBTITLE

An Investigation into the Application of Artificial Neural Networks to Adaptive Optical Imaging Systems

5. FUNDING NUMBERS

6. AUTHOR(S)

Andrew H Suzuki, Capt, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GEO/ENG/91D-04

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Capt Michael Roggemann
Phillips Lab/ARCI
Kirtland AFB, NM 87117

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Recurrent and feedforward artificial neural networks are developed as wavefront reconstructors. The recurrent neural network studied is the Hopfield neural network and the feedforward neural network studied is the single layer perceptron artificial neural network. The recurrent artificial neural network input features are the wavefront sensor slope outputs and neighboring actuator feedback commands. The feedforward artificial neural network input features are just the wavefront sensor slope outputs. Both artificial neural networks use their inputs to calculate deformable mirror actuator commands. The effects of training are examined.

14. SUBJECT TERMS

Artificial Neural Network, Adaptive Optics

15. NUMBER OF PAGES
86

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to **stay within the lines to meet optical scanning requirements.**

Block 1. Agency Use Only (Leave Blank)

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Names(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency. Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of ..., To be published in .... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availablity Statement. Denote public availability or limitation. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR)

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - DOD - Leave blank
DOE - DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports
NASA - NASA - Leave blank
NTIS - NTIS - Leave blank.

Block 13. Abstract. Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS only).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

AFIT/GEO/ENG/91D-04

AN INVESTIGATION OF THE APPLICATION OF ARTIFICIAL NEURAL NETWORKS TO ADAPTIVE OPTICS IMAGING SYSTEMS

THESIS

Presented to the Faculty of the School of Engineering

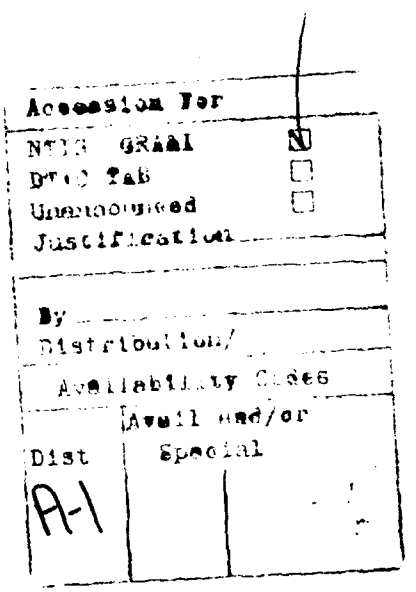
of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Electrical Engineering



Andrew H. Suzuki, B.S.E.E

Captain, USAF

December 1991

Approved for public release; distribution unlimited

Preface

This thesis develops a baseline for further research at AFIT into the topic of integrating artificial neural networks into the adaptive optical imaging system environment. It is my sincere hope that I will make the work of future AFIT Master's students a little easier.

I am deeply indebted to my faculty advisor, Capt Byron M. Welsh, for his continuing patience and assistance throughout these troubled times. Capt Welsh's persistence and seemingly unending knowledge served as an inspiration and guiding light in the darkness I sometimes viewed as the thesis black hole.

I also wish to thank my thesis sponsor Capt Michael Roggemann and the my thesis committee members, Maj. Steve Rogers, Dr Matthew Kabrisky, and Capt Dennis Ruck, for their infinite wisdom and assistance.

Finally, the most deserved thanks goes to my wife Roni and my children Anthony, Matthew, Melissa, and Michael for their understanding, patience, and support throughout the "AFIT experience."

Andrew H. Suzuki

Table of Contents

List of Figures

List of Tables

Abstract

This thesis develops artificial neural network wavefront reconstructors for a ground-based adaptive optical imaging system. The incoming object wavefront is reflected from a 21 actuator continuous faceplate deformable mirror onto a 20 subaperture Hartmann-Shack wavefront sensor.

Recurrent and feedforward artificial neural networks are developed as wavefront reconstructors. The recurrent neural network studied is the Hopfield neural network and the feedforward neural network studied is the single layer perceptron artificial neural network. The recurrent artificial neural network input features are the wavefront sensor slope outputs and neighboring actuator feedback commands. The feedforward artificial neural network input features are just the wavefront sensor slope outputs. Both artificial neural networks use their inputs to calculate deformable mirror actuator commands. The effects of training are examined for the single layer perceptron neural network.

The adaptive optical system is simulated by integrating simulation software provided by the Phillips Laboratory Kirtland AFB, NM and a computer model of the neural network controller. The simulation code contains adaptive optical system models and an atmospheric phase screen generator.

This thesis research supports the feasibility of integration of artificial neural networks with adaptive optical imaging systems. The simulation results show the performance of artificial neural network reconstructors are commensurate with conventional linear estimation techniques. The simulation also shows, that training artificial neural networks can provide performance enhancements not currently achievable using conventional techniques.

AN INVESTIGATION OF THE APPLICATION OF ARTIFICIAL NEURAL NETWORKS TO ADAPTIVE OPTICAL IMAGING SYSTEMS

I. Introduction

1.1 Motivation

The traditional Air Force mission, "to fly and fight" is becoming increasingly dependent on peripheral technologies. High resolution ground based exo-atmospheric imaging technology is crucial to successful fulfillment of present and future Air Force mission profiles. The success of high resolution ground based exo-atmospheric imaging hinges on the solution to one crucial problem, negating the effects of atmospheric turbulence.

Atmospheric turbulence significantly degrades the performance of exo-atmospheric imaging systems. Atmospheric turbulence produces image motion, distortions, and intensity fluctuations (scintillation) (16:360). There are several methods to compensate for turbulence induced image degradation. For ground based applications, the most common methods are post processing and predetection. Post processing compensates for image aberrations after image retrieval. Predetection compensates for image aberrations in real-time. This thesis addresses only predetection atmospheric imaging, specifically the use of adaptive optics.

Predetection or adaptive optics uses real-time control to compensate for image degradations encountered when imaging through turbulence. Real time imaging is essential to the exo-atmospheric tracking and possibly weapons delivery aspects of the Air Force's mission profile.

The adaptive optics system performs three basic functions: wavefront sensing, reconstruction, and compensation. Wavefront sensing measures the turbulence induced phase aberrations of the incoming wavefront across the system aperture. Wavefront reconstruction uses the sensed phase aberrations to calculate the deformable mirror actuator commands. Wavefront compensation receives

the actuator commands and positions individually controlled actuators on a deformable mirror to counteract the effects of turbulence.

The integration of the artificial neural network with the adaptive optical imaging system occurs in the wavefront reconstruction process. This thesis research develops "artificial neural network reconstructors." The artificial neural network reconstructors use slope sensor measurements to calculate deformable mirror actuator commands.

The Air Force's Phillips Laboratory is at the forefront of the DOD's effort to explore adaptive optics as a means to compensate for the effects of atmospheric degradation on exo-atmospheric imaging systems. The Phillips Laboratory, who sponsors this thesis, is actively investigating the use of adaptive optical systems to image exo-atmospheric objects, such as satellites.

1.2 Approach

The goal of this thesis is to demonstrate that artificial neural networks can accomplish adaptive optical wavefront reconstruction. Artificial neural networks have many features that can potentially benefit present and future adaptive optical imaging systems. The artificial neural network combines the features of a highly parallel architecture and extremely fast computational speed to provide significant advantages over conventional wavefront reconstruction techniques. The added feature of artificial neural network training can provide accurate system output estimates under a wide variety of system operating environments. To accomplish the research goal, the following tasks must be completed:

1. Determine the types of artificial neural networks to use as research baselines.
2. Develop artificial neural network wavefront compensation algorithms using conventional adaptive optics control techniques.
3. Implement adaptive optical wavefront compensation using the artificial neural network compensation algorithms.
4. Simulate and evaluate the system performance by comparing with conventional linear estimation techniques.

5. Demonstrate the potential benefits of artificial neural networks.

The first task starts with a delineation between the types or classes of artificial neural networks that are candidate wavefront reconstructors. The two classes of neural networks researched are the feedforward and recurrent artificial neural network. This thesis analyzes a representative artificial neural network from each class. The baseline artificial neural networks chosen were the; Hopfield neural network for the recurrent class and the single layer perceptron for the feedforward class.

The second task maps the adaptive optics wavefront reconstruction process into each neural network state space. Recurrent and feedforward neural networks have unique properties which govern their implementations as wavefront reconstructors. Each neural network is studied to determine the best means of integration with the adaptive optical control system.

The third task develops software models of the artificial neural network wavefront reconstructor. The artificial neural network models are integrated with a software simulation package provided by the Phillips Laboratory, Kirtland AFB, NM. The Phillips Laboratory simulation package contains the adaptive optical system models and control algorithms for conventional adaptive optics compensation. Modifications are made to the simulation code to replace the conventional reconstructor with the artificial neural network reconstructor.

For the fourth task, an unmodified version of the simulation code is the comparison baseline. The unmodified version uses a conventional linear estimation technique known as the "least squares reconstructor." System performance comparisons are made between the results from the least squares reconstructor and the artificial neural network reconstructor.

The final task demonstrates the potential benefits of artificial neural network compensation. The feature of artificial neural network training is used to show how artificial neural networks can provide performance enhancements not currently achievable using conventional reconstruction techniques.

1.3 Scope

This thesis is the initial step into the integration of artificial neural networks and adaptive optical imaging at AFIT. Several estimates and assumptions are used to expedite progress and complete the research tasks in the allotted time. Most of the estimates and assumptions have their roots based in precedents found in the literature. Some of the assumptions are imposed by the software simulation code and others are pure engineering judgement.

This thesis assumes that the adaptive optical imaging system is a ground based system. A separate system tracks the object and its associated reference or guide star. The object and guide star are in close proximity to one another, therefore adaptive optic compensation for the guide star also compensates for the object. A separate imaging system removes the overall piston and tilt of the incoming wavefront.

The simulation code produces a random phase screen developed from a technique introduced by Cochran (4). This phase screen simulation produces phase screens whose structure functions, when computed over a large ensemble, corresponds to a Kolmogorov atmosphere (13:389-391). This simulation code also contains models for each component of the adaptive optical imaging system. The wavefront sensor model is patterned after the Hartmann-Shack sensor (HS WFS), which produces slope outputs. The detection process produces random additive noise which corrupts the slope outputs. The noise and signal are uncorrelated. The deformable mirror is modeled as a continuous face sheet mirror with Gaussian actuator influence functions. This thesis neglects the effects of a time delay between the wavefront sensing and correction. The adaptive optical system geometry used for this thesis has 20 wavefront sensor subapertures and 21 deformable mirror actuators (Figs 1.1, 1.2). This particular geometry is an artifact of the method used by the simulation code to determine the actuator and subaperture locations. The operational parameters used in this thesis are shown in Table I.

The Hopfield neural network is modeled (Fig. 1.3) as a discrete time system with synchronous updates (32). This means the updates to the Hopfield outputs and inputs are made at the same time and time is modeled as discrete steps. The neuronal outputs of the Hopfield network correspond to

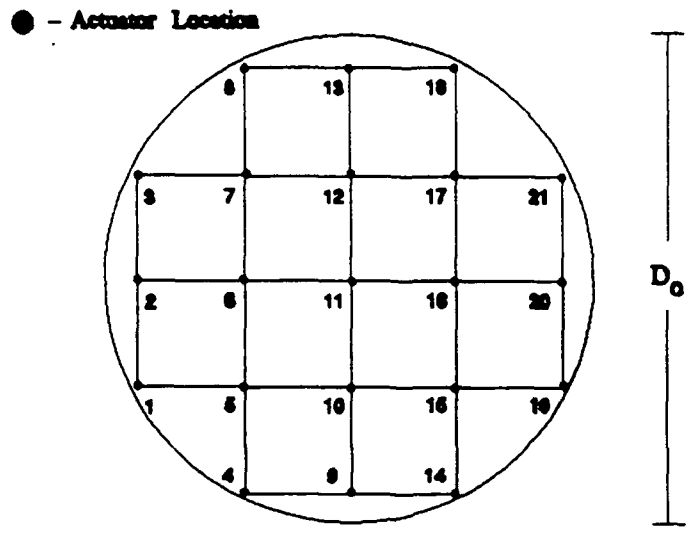


Figure 1.1 Deformable Mirror Actuator Geometry

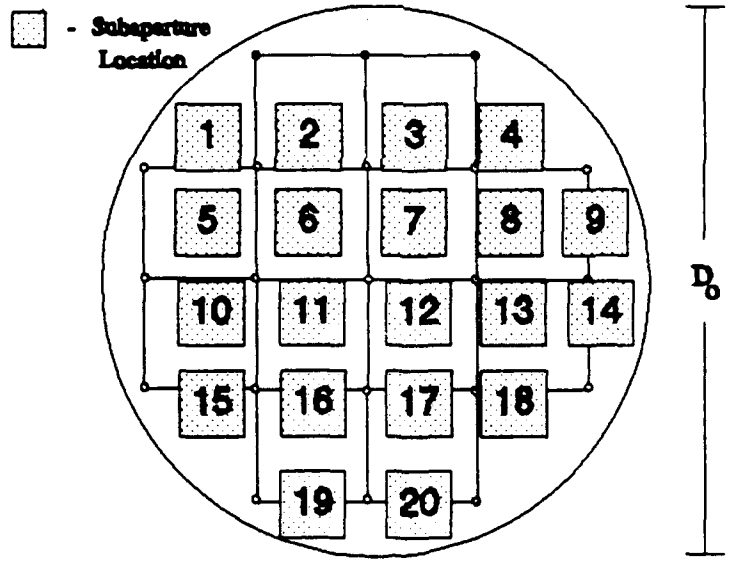


Figure 1.2 Hartmann Wavefront Sensor Subaperture Geometry

the individual actuator commands to the deformable mirror. The input features to a neuron are the weighted slopes from the wavefront sensor and neighboring actuator feedback commands (Fig. 1.4). To produce an output command, the weighted inputs are summed and passed through a soft limiter activation function.

Table I System Parameters

Parameter	Value
Wavelength (λ)	500 nm
Actuator Separation (actsep)	15 cm
Mirror Diameter (D_o)	75 cm
Light Level (N_p)	1-1000 photoevents/subap/ integration time
Atmospheric Coherence Diameter (r_o)	7.5, 10, 20 cm

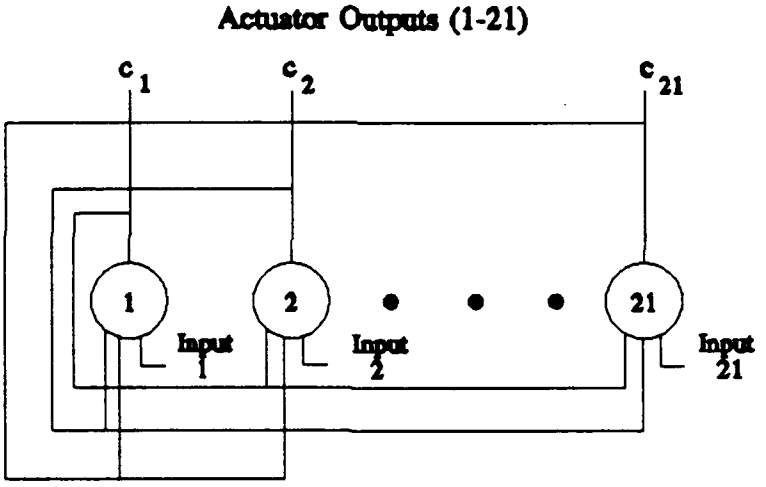


Figure 1.3 Hopfield Neural Network

The neuronal outputs of the single layer perceptron (Fig. 1.5) are the actuator commands to the deformable mirror. For the single layer perceptron, the neuronal inputs are the weighted slope sensor outputs (Fig. 1.6). The output nonlinearity or activation function used is a sigmoid or "squashing function."

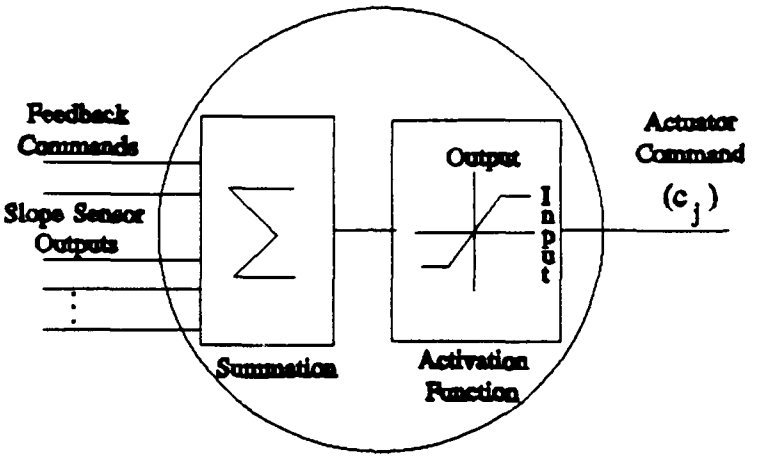


Figure 1.4 Hopfield Neuron

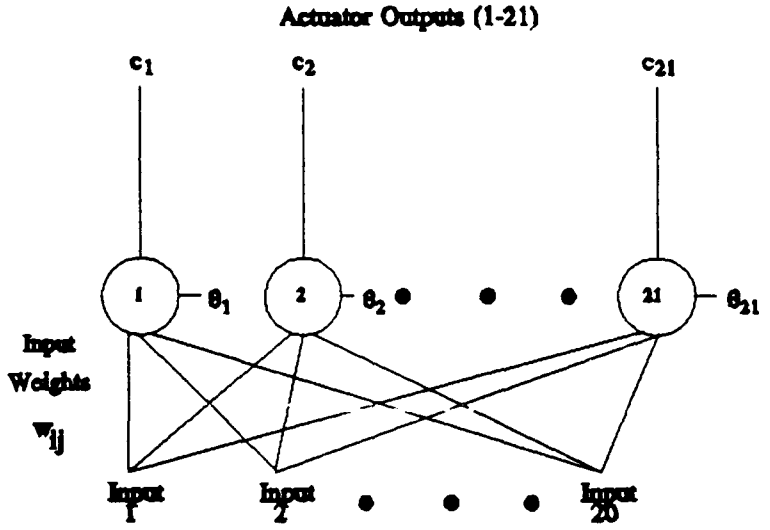


Figure 1.5 Single Layer Perceptron Artificial Neural Network

1.4 Standards

This thesis analyzes the system performance in response to variations of both system and hardware parameters of the adaptive optical imaging system. The system parameters are the light level and seeing conditions. The hardware parameter is the amount of interconnections from the wavefront sensor to the wavefront reconstructor.

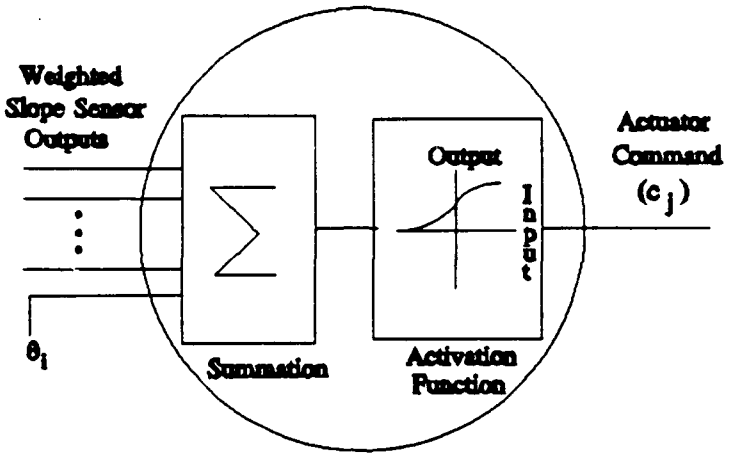


Figure 1.6 Single Layer Perceptron Neuron

As a final product, this thesis contrasts the performance between the conventional least squares reconstruction algorithm and the artificial neural network reconstructors. The primary figure of merit used is the mean square residual phase error between the estimated and the actual wavefront.

1.5 Chapter Outlines

The following is a brief synopsis of the information found in each of the thesis chapters.

1.5.1 Chapter 2. Chapter 2 presents the fundamental concepts behind adaptive optical imaging systems and artificial neural networks. These fundamental concepts are essential to understanding the problems that plague the adaptive optical imaging system and a possible somution using artificial neural networks. The concepts presented in this chapter serve as the basis for the algorithms and analysis presented in chapter 3.

1.5.2 Chapter 3. This chapter uses the background knowledge provided in chapter 2 to develop feedforward and recurrent reconstructor models. Each reconstructor model is then validated as being either a Hopfield or a single layer perceptron artificial neural network.

1.5.3 Chapter 4. This chapter presents the results of the analysis and simulation. The task of analyzing and simulating the adaptive optical imaging system is an intricate process. To adequately address each specific part, the entire simulation process was divided into specific experiments. Each experiment and the results are discussed in detail.

1.5.4 Chapter 5. This chapter states the conclusions based upon the results provided in Chapter 4. The chapter concludes with recommendations for areas of future study.

1.6 Summary

The adaptive optics imaging system presents a very complex control system problem. A proposed solution is the use of an artificial neural network wavefront reconstructor. This thesis investigates the application of two different types of artificial neural networks (i.e. Hopfield and single layer perceptron artificial neural network) to the adaptive optical imaging system. The artificial neural network reconstructor performance and potential performance enhancements using artificial neural network training are presented.

II. Background

2.1 Introduction

This chapter presents the fundamental concepts behind adaptive optical imaging systems and artificial neural networks. These fundamental concepts are essential to understanding the problems that plague the adaptive optical imaging system and a possible solution using artificial neural networks. The concepts presented in thesis chapter serve as the basis for the algorithms and analysis presented in the next chapter.

This chapter provides the background concepts needed to obtain a fundamental understanding of the problem (atmospheric turbulence), system (adaptive optical imaging system), and proposed solution (artificial neural networks). Prior to developing artificial neural network reconstructors, the following research questions must be answered:

1) What is the origin of the problem ?
2) How does the adaptive optical imaging system work ?
3) How does the problem influence the system ?
4) Is there an alternate solution to the problem ?
5) How does the solution work ?
6) Has the solution been implemented ?
7) How will this research effort extend the present implementation ?

This chapter sequentially addresses these fundamental questions. Research question 1 is answered in section 2.2. Section 2.2 provides an overview of the concepts of the most important problem facing present and future optical imaging systems; negating the effects of atmospheric turbulence (research question 1). Section 2.2 discusses why phase only correction is used to combat the effects of atmospheric turbulence. Section 2.2 also introduces a useful measure of the image degradation effects introduced by atmospheric turbulence, the atmospheric coherence diameter (r_o). Section 2.3 answers research question 2 by defining three basic functional elements (sensing, reconstruction, and compensation) of the adaptive optical imaging system and how these functional elements are currently

implemented. The system limitations imposed by atmospheric turbulence are presented in section 2.4 (research question 3). Research question 4 is addressed in section 2.5 by presenting a brief discussion of how artificial neural networks represent an alternate solution to the problem. The theory behind the solution (artificial neural networks) is presented in section 2.6 (research question 5). Research question 6 is answered in section 2.7 by discussing the current implementations of the artificial neural network solution. This chapter concludes by presenting the direction that this research effort will take to address the final research question.

2.2 Turbulence

This section discusses the origin of the problem (research question 1) facing adaptive optical imaging system. The origin of the problem is atmospheric turbulence. Atmospheric turbulence

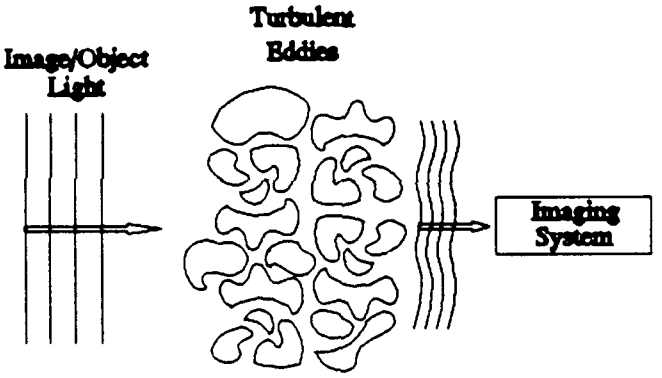


Figure 2.1 Turbulent Eddy Representation of the Atmosphere

degrades the performance of optical imaging systems. The aberrations produced by the atmosphere are caused by inhomogeneities of the refractive index of air. Differential heating of the Earths' surface causes large scale temperature inhomogeneities. These large scale inhomogeneities are broken up into smaller scale "eddies" by convection and turbulent wind flow. Each turbulent eddy has its own individual refractive index (13:388). The localized refractive indices temporally and spatially modulate the amplitude and phase of a propagating wavefront. Figure 2.1 graphically displays the effect of the turbulent eddies on plane wave propagation. The plane wave propagating through the atmosphere

is sensed as an aberrated wave by the imaging system. The aberrations are due to amplitude and phase modulations induced by atmospheric turbulence.

The amplitude modulation effects are known as scintillation. An example of scintillation is the twinkling of stars (34:224). The scintillation effects can be considered negligible when the following condition is satisfied (36:2580)

$$\sqrt{L\lambda} \ll D_o \tag{2.1}$$

where L is the objects distance from the exit pupil, D_o is the pupil diameter, and λ is the wavelength. Although this condition is not strictly met at high altitudes, much of the literature suggests that the amplitude modulation effects are less severe than the phase modulation effects for ground based vertical imaging through the atmosphere (9) (11) (25) (38). Therefore, by only compensating for this phase aberration, the object image can, in theory be reconstructed. This research effort focusses only on the phase compensation aspect of adaptive optical systems.

2.2.1 Atmospheric Coherence Diameter. The atmospheric coherence diameter or Fried parameter is a useful measure of the image degradation caused by atmospheric turbulence. The atmospheric coherence diameter is a measure of the correlation qualities of the optical wave propagating through the atmosphere. Typical values of r_o at a good mountain top observatory can range from 5 cm to 20 cm (13:432).

The atmospheric coherence diameter is defined by (13)

$$r_o = 0.185 \left(\frac{\lambda^2}{\sec(\zeta) \int C_n^2(z)\, dz} \right)^{3/5} \tag{2.2}$$

where λ is the wavelength, ζ is the zenith angle, $C_n^2(z)$ is the atmospheric structure constant, and z is a point along the optical path.

Subsequent sections will show how the atmospheric coherence diameter is intimately related to the image quality and resolution of the adaptive optical imaging system. Section 2.2.2 shows the expressions for the atmospheric optical transfer function are easier to understand and simpler in form when put in terms of the atmospheric coherence diameter. Section 2.2.3 shows the resolution of an adaptive optical imaging system is directly related to the ratio of the systems aperture diameter and the atmospheric coherence diameter.

2.2.2 Image Quality. A performance measure of an imaging system operating in the presence of the atmosphere is the average optical transfer function (OTF) (28). The average OTF is used because the turbulence is a random process in time and space and the ensemble or average quantities are of interest. Comparison between the system optical transfer function and the diffraction limited transfer function gives an excellent measure of the system performance and image quality.

The Rytov solution for the propagation of an electromagnetic wave in a homogeneous, isotropic medium with zero mean Gaussian random phase statistics defines the long exposure OTF (13). The long exposure OTF of the medium is a function of the atmospheric phase structure function. The expression for the phase structure function, D_ϕ, within the inertial subrange is given by (13)

$$D_\phi(r) = 6.88 \left(\frac{r}{r_o}\right)^{5/3} \tag{2.3}$$

where r is the spatial separation between two points on the object wavefront and r_o is the atmospheric coherence diameter. The long exposure OTF of the atmosphere can be shown to be

$$H_{avg}(r) = \exp\left[-\frac{1}{2} D_\phi(r)\right] \tag{2.4}$$

The overall OTF of the optical system and the atmosphere is the product of the OTF of the optical system and the average OTF of the atmosphere. The overall OTF and it Fourier transform or the

point spread function, defines the image quality of the adaptive optical imaging system. Equations 2.2 and 2.3 exemplify the dependence of image quality on the atmospheric coherence diameter.

2.2.3 Resolution. The previous section pointed out the effect of the atmospheric coherence diameter on the imaging qualities of the adaptive optical imaging system. This section extends the analysis of the atmospheric coherence diameter by examining its effects on the system's resolution ($\Re$).

The diffraction limited resolution ($\Re_{max}$) between two point sources is represented by the relationship (12:130)

$$\Re_{max} = 1.22 \frac{\lambda d_i}{D_o} \tag{2.5}$$

where d_i is the distance from the aperture to the image plane. Turbulence effects significantly reduces the resolution from that given in equation 2.5. A useful measure of the resolution attainable is based upon the coherence diameter to aperture diameter ratio.

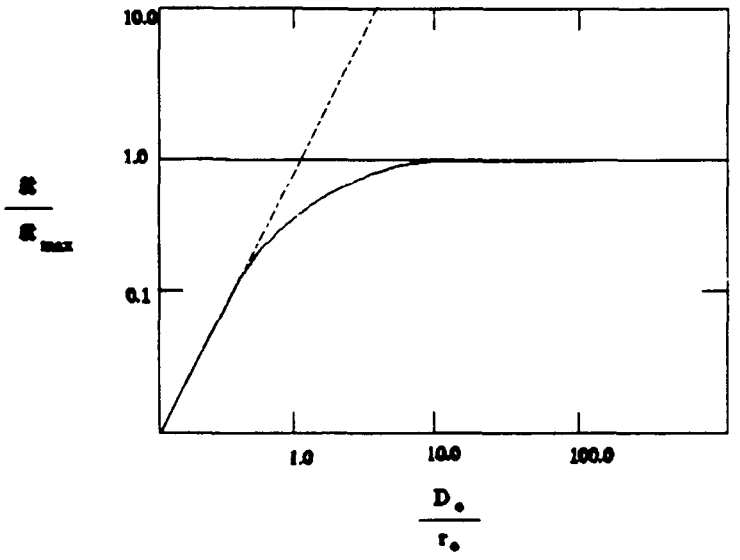


Figure 2.2 Normalized Resolution vs Normalized Diameter

Figure 2.2 (13:431) shows a plot of the normalized resolution of an imaging system versus the normalized aperture diameter. $\Re$ and $\Re_{max}$ are the resolution and the maximum resolution attainable

by the system. D_o is the aperture diameter and r_o is the atmospheric coherence diameter. The dashed line represents a diffraction limited imaging system, and the solid line represents a system subjected to atmospheric turbulence effects. As figure 2.2 shows, the resolution of systems subjected to atmospheric turbulence effects is significantly altered. When the aperture diameter is larger than r_o, the maximum resolution achievable by the imaging system is no longer a function of the aperture diameter. Instead, the maximum resolution is limited by the ratio of the aperture diameter to the atmospheric coherence diameter. As the aperture diameter to r_o ratio increases beyond unity, larger aperture diameters offer no significant improvement in resolution. The resolution of larger aperture systems coincides with the resolution of a system that has an aperture diameter equal to r_o.

To exemplify this phenomena, under identical operating conditions, the maximum resolution of a 3 m telescope subjected to atmospheric turbulence is comparable a same telescope with an aperture equal to r_o (5 to 20 cm).

2.3 Adaptive Optical Imaging Systems

The primary function of an adaptive optics imaging system is to provide real-time imagery of exo-atmospheric objects, such as satellites. The generic adaptive optics imaging system (Fig 2.3) must

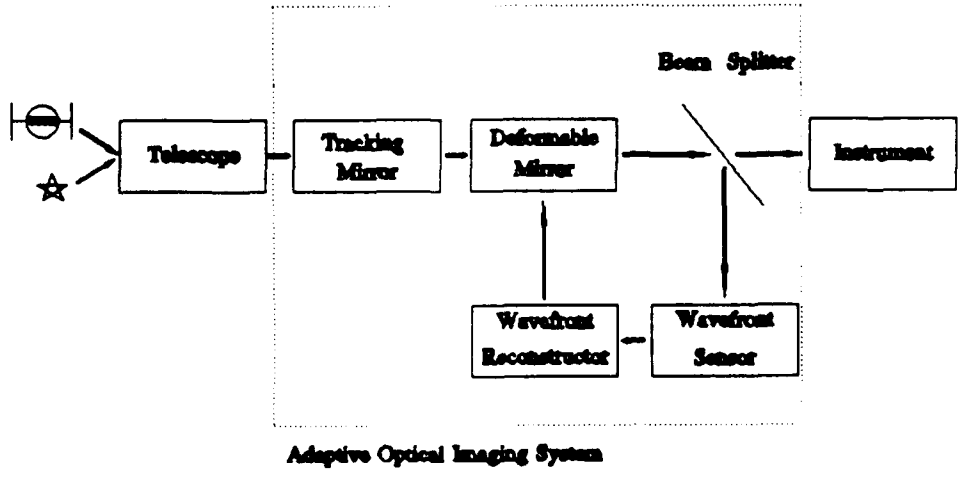


Figure 2.3 Generic Adaptive Optical Imaging System

accomplish three basic functions to negate the effects of atmospheric turbulence. The three functions are to sense, reconstruct, and compensate for phase distortions induced by atmospheric turbulence. This section answers research question 2 by briefly discussing the three basic functions of the adaptive optical imaging system.

2.3.1 Wavefront Sensing. Assuming the overall wavefront tilt is removed, the wavefront sensor now measures the residual phase gradients present in the incoming wavefront. A wavefront sensor senses the deviation of the incoming wavefront from an ideal plane wave. The wavefront sensor samples the incoming wavefront at intervals known as "subapertures" (Fig 1.2). Each subaperture senses the localized slope deviation of the incoming wavefront. The two most common wavefront sensors are the shearing interferometer and the Hartmann-Shack wavefront sensor (Fig 2.4). A model

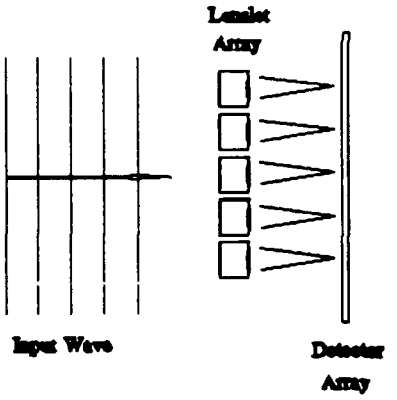


Figure 2.4 Hartmann-Shack Wavefront Sensor

of the Hartmann-Shack wavefront sensor is used for this thesis.

2.3.2 Wavefront Reconstruction. Wavefront reconstruction receives the slope sensor information and calculates the actuator commands needed to negate the phase aberrations induced by atmospheric turbulence. The wavefront sensing process gives slope sensor measurements that are a linear function of the turbulence induced phase aberration. Due to this linear relationship between

the slope sensor measurements and the phase aberrations, conventional linear estimation techniques are used to reconstruct the object wavefront (6:1). The two most common linear estimation techniques are termed the least squares and the minimum variance reconstructors (28:1). This thesis uses the least squares reconstructor as the comparison baseline.

2.3.3 Wavefront Compensation. The deformable mirror receives actuator commands from the wavefront reconstructor to provide a phase conjugate of the wavefront phase aberrations. The mirror uses individually controlled actuators to make a piecewise approximation of the phase conjugate wavefront surface. The 2 most common types of deformable mirror architectures are the segmented mirror and continuous face sheet mirror (Fig 2.5).

The segmented mirror has been used for atmospheric optical compensation for many years. Segmented mirrors can compensate for piston component of the aberrated wave or piston and tilt. In a segmented mirror, the actuator phase compensation at one segment is not affected by actuator control of any other segment. Early integration of adaptive optical imaging and artificial neural networks used segmented mirrors (31:19).

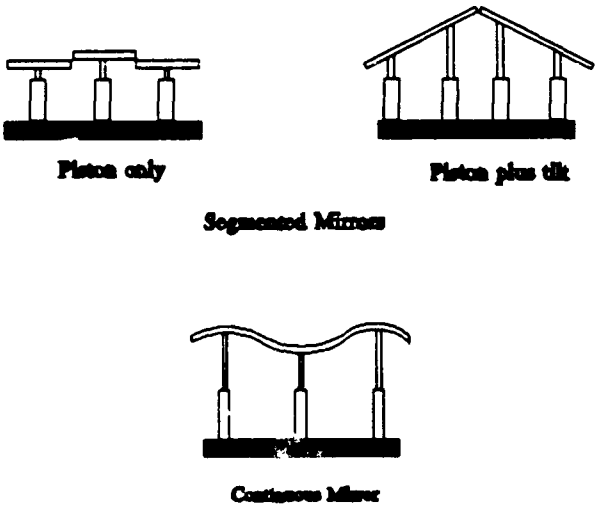


Figure 2.5 Deformable Mirror Geometries

Continuous facesheet mirrors use a thin quartz or glass facesheet supported on arrays of discrete actuators. These discrete actuators deform or put a "dimple" in the mirror surface to provide a conjugate phase correction. This thesis uses a continuous facesheet mirror as a baseline. In a continuous facesheet mirror, any given actuator command influences the mirror surface over the entire mirror diameter. This thesis models this effect by using Gaussian actuator influence functions. The actuator influence functions are defined at each actuator center and any phase compensation point is found to be a combination of all the actuator influences summed at that point.

2.4 Fundamental Limitations on the Adaptive Optical Imaging System

The atmosphere imposes fundamental limitations on the adaptive optical imaging system (research question 3). These limitations can be loosely grouped as limitations on; the computational frame time of the adaptive optical imaging system, the image quality, resolution, and the light level. This section presents a discussion of the limitations imposed on the adaptive optical imaging system by atmospheric turbulence.

2.4.1 Computational Frame Time. The turbulent nature of the atmosphere produces wavefront phase aberrations which exhibit a high degree of temporal instability. This temporal instability imposes a fundamental limitation on the computational frame time for adaptive optical compensation. The adaptive optics system must sense and compensate for the wavefront aberrations within the atmosphere's coherence time (τ_c). The atmospheres coherence time is the time delay that a wavefront and a temporally delayed version of the wavefront can undergo and still be highly correlated (i.e. temporally stable) (13:158). The adaptive optics system must have a closed loop control bandwidth of at least 100 Hz to guarantee temporal stability of the object wavefront. This closed loop bandwidth results in a step response settling time of less than 10 milliseconds (7:260) (8:224). The millisecond frame rate requirement for real-time operation coupled with the large number of control channels required for adaptive optical imaging systems imposes severe restrictions for current and future adaptive optics imaging systems.

The conventional wavefront reconstruction process is a computationally intense environment. The number of active control channels and the computational frame time limitations make the reconstruction process difficult to implement without sophisticated and extremely fast hardware and software. Coupling the magnitude of the control problem with the variational nature of factors external to the adaptive optics imaging system, make the task of exo-atmospheric imaging quite formidable.

2.4.2 Light Level. The photon count per subaperture or "light level" is another fundamental limitation on the adaptive optical imaging system. The detectability of the wavefront phase aberrations is directly related to the amount of light emanating from the object. At very low light levels, shot noise effects in the wavefront sensor can make attempts to extract an image virtually impossible. The wavefront sensor outputs are modeled as a signal plus additive noise measurement. The additive noise or "slope measurement noise" is due to the shot noise effects of the detection process and is assumed to be uncorrelated with the wavefront phase. The noise is modeled as a zero mean Gaussian random process, where the standard deviation of the slope measurement noise σ_ϵ is defined by

$$\begin{aligned} \sigma_\epsilon &= \frac{0.86\,\pi\,\eta}{\sqrt{N_p}\,r_o} && D_o \geq r_o \\ &= \frac{0.74\,\pi\,\eta}{\sqrt{N_p}\,D_o} && D_o < r_o \end{aligned} \tag{2.6}$$

where η is a correction factor which compensates for the finite sized detectors separated by unusable space and N_p is the total subaperture photon count. The value of η is assumed to be 1.5. This value corresponds to typical values found for charge coupled device technology (38:1919). Equation 2.6 points out the inverse relationship of the additive noise to the light level. As the light level increases the variance of the slope measurement noise decreases.

2.5 *An Alternate Solution*

Research question 4 asks "Is there an alternate solution to the problem?' Recently, Lockheed presented an alternate solution to alleviate some of the speed and interconnection problems encountered with conventional analog and digital wavefront reconstructors (7:260-268) (8:222-229). This solution integrated artificial neural networks into the wavefront reconstruction process. The data rate of the artificial neural network reconstructor (5000 Hz) is almost an order of magnitude faster than a conventional STAR array processor compensation method (600 Hz). Artificial neural networks can also reduce the complex interconnection schemes encountered with larger deformable mirror and wavefront sensor arrays. Although this integration of artificial neural networks and adaptive optics is still in the early stages of development, the preliminary results are quite promising.

2.6 *Artificial Neural Networks*

This section discusses the operational characteristics of artificial neural networks (research question 5). The generic artificial neural network is a highly parallel, distributed processing structure that consists of simple processing elements. Each processing element, or neuron has a single output that may branch out to many peripheral connections. The output signal can have any functional form, but must depend on the current values at the input of the neuron. Neural networks have many names, such as connectionist models, distributed processing models, or neuromorphic systems (23:4). Whatever the nomenclature, all of these models attempt to provide performance enhancements in a calculation intense environment by using densely packed, highly parallel processing elements. Two of the basic characteristics used to describe artificial neural networks are the type of network (feedforward/recurrent) and the learning technique (21:111-221) (27:39-40).

2.6.1 Recurrent and Feedforward Neural Networks. Feedforward neural networks (Fig. 2.6) operate on the sum of weighted inputs using an activation function or nonlinearity to calculate an output. For single layer feedforward networks (Fig. 1.5) the activation function outputs are the system outputs. For multiple layer feedforward neural networks (Fig. 2.6) the activation function outputs

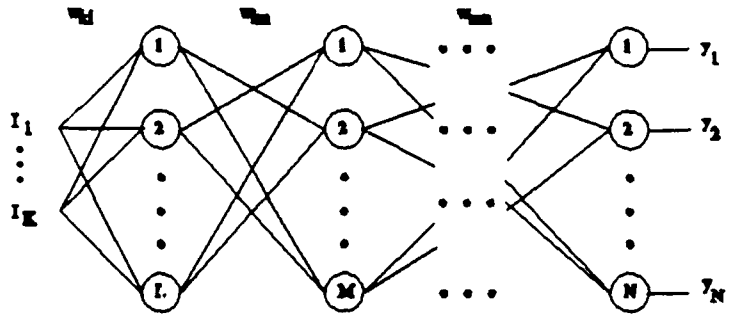


Figure 2.6 Multilayer Feedforward Artificial Neural Network

become weighted inputs to the next layer of neurons, and so on, until the output layer produces the system output. Outputs in this type of neural network are changed only when a new input exemplar is presented. The weights in these type of networks are often trained to produce desired outputs. The feedforward artificial neural network studied for this thesis is the single layer perceptron.

A recurrent neural network (Fig. 2.7) is a neural network that has feedback paths from neuronal outputs to neuronal inputs. The output response of these networks is dynamic. The outputs continually changing until the network converges to reach a stable value. An input is presented to the network then an output is calculated. The outputs are then fed back into the network as inputs. This process is repeated until a stable condition is reached. Stability can be defined as convergence to a single output, as in the case of a Hopfield neural network. Convergence can also be defined as a predetermined time in the future, as in the case of a recurrent backpropagation neural network. The recurrent artificial neural network examined for this thesis are the Hopfield neural network, and an extension of the Hopfield network that takes advantage of the learning properties of artificial neural networks.

2.6.2 Artificial Neural Network Learning. Artificial neural networks can be trained to learn relationships between patterns/exemplars. Learning is the encoding of information, a system can learn

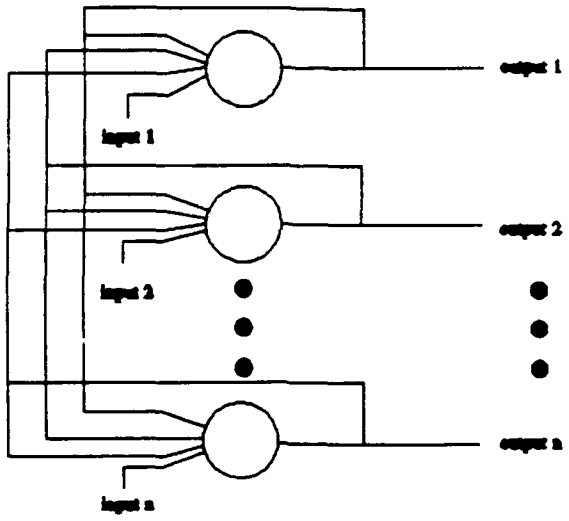


Figure 2.7 Recurrent Artificial Neural Network

a pattern by encoding information about that pattern. Subjecting a neural network to many patterns allows the network to generalize about the interrelationship between the patterns. These interrelationships are formed by statistical process inference or class relationship between the ensemble of patterns. The ways in which artificial neural networks learn are described as being either unsupervised, supervised, or a combination of both.

In unsupervised training the training feature data set is given to the network. The data along with the training algorithm position the weights into clusters that reflect the distribution of the training data (40:2-9). The unsupervised training algorithm uses the data to approximate the probability distribution of the ensemble data set. The inference of the probability distribution is used to classify subsequent test data.

Supervised learning presents the neural network with an ensemble of training data pairs composed of input data features and their desired output. The figure of merit is the difference between the neural network calculated output and the desired output. The interconnection weights are then trained to minimize some function of this error. Typically, the weights are trained to reduce the mean squared error between the calculated and desired outputs. The training process is repeated over many training pairs to allow the network to learn the statistical relationship binding the entire

data set. This thesis addresses a specific supervised training technique known as the backpropagation method (17:593) (39:270-280). The backpropagation training method will be applied to the single layer perceptron and the Hopfield artificial neural networks.

2.6.3 The Hopfield Neural Network. The Hopfield neural network or relaxation neural network is the network of choice for many hardware implementations. The reasons are the conventional Hopfield neural network has fixed interconnection weights (27:78-82) and extremely fast convergence properties. For these reasons, the Hopfield neural network is the baseline recurrent neural network used for this thesis.

The Hopfield neural network has been applied to many classes of problems. These classes can be loosely classified by the type of problem the neural network is attempting to solve. These two application groups are the associative memory (1:204) and function estimation (10:805-810) (41:499-506) (33:533-541) applications. This thesis diverges from the associative memory applications and applies the Hopfield neural network to a classical optimization problem, adaptive optical wavefront reconstruction. The operational concept of the Hopfield neural network begins by presentation of an input to the network. The network continually operates on the weighted sum of inputs and feedback outputs, until the network converges to a stable state. The Hopfield neural network can processes the weighted sums by passing them through a nonlinear function such as a hard limiter or sigmoid (squashing function). This thesis uses a soft limiter or ramp nonlinear function to calculate neuronal outputs. When using Hopfield neural networks, fundamental issues must be addressed to adequately predict and characterize the systems performance. The most important issues are to a Hopfield neural network designer are convergence stability and local minima (Fig 2.8).

Convergence stability posed a formidable problem to early researchers. It was very hard to predict which networks would be stable. This problem continued until fundamental work was accomplished by Cohen and Grossberg to characterize the convergence stability for this specific class of recurrent neural networks. The dynamic Hopfield neural network system possesses convergence stability if there is a Lyapunov function that describes the performance. The existence of a Lyapunov

function proves stability, whereas the nonexistence of the Lyapunov function does not infer anything about system stability (21:69-73). The Lyapunov function of the Hopfield neural network is often termed the "energy" function (18:3089-3090). Examination of the gradient of the energy function with respect to the system outputs can prove convergence stability. The energy function of the Hopfield neural network is intimately dependent upon the properties of the feedback weight interconnection matrix. Exhaustive analysis of the convergence issues has resulted in two requirements for the feedback interconnection weight matrix (5:815-826). The first is there is no self feedback for the Hopfield neurons, or the weight from the i^{th} to the i^{th} neuronal element is zero. The final restriction imposed by convergence stability is that the feedback interconnection weight matrix must be symmetric about the main diagonal, or the weight from the i^{th} to the j^{th} neuronal element equals the weight from the j^{th} to the i^{th} neuronal element.

When an energy surface becomes complicated and has many peaks and valleys, local minimas

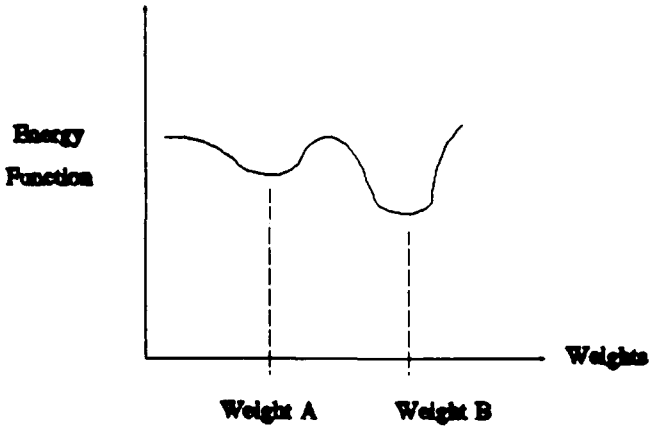


Figure 2.8 Local Minimum

become a significant concern (22:137) (30:327). Issues of numerous local minima plagued early researchers using Hopfield neural networks. Local and global minimas represent stable solutions in the energy state space for this class of neural networks. When a Hopfield neural network converges,

it settles into a minimum on the energy surface. Convergence to a local minima represents an incorrect solution.

2.6.4 Single Layer Perceptrons. The single layer perceptron artificial neural network was originally proposed by Rosenblatt in 1959 (27:47). The single layer perceptron artificial neural network and a single perceptron are shown in Figs. 1.5 and 1.6. The inputs to the perceptron can be from other perceptrons or the external sources (e.g. slope sensor outputs). Each input is multiplied by a synaptic weight w_{ij} and summed. A bias term (θ_i) is added and a nonlinear transformation is applied. There are several forms for the nonlinearity, for example a hard limiter, a linear ramp, or a sigmoid. The bias term can also be treated as another weight multiplied by a constant of one. The single layer perceptron uses the weights to map the vector of inputs onto a vector of outputs.

The key to function estimation using perceptrons is using a learning algorithm to find the correct set of weights. The most common perceptron learning algorithm is called "backward error propagation," or simply "backprop." Backprop calculates the error at the output layer and backpropagates the error through the network correcting the weights. The correction represents a gradient descent down the error surface toward a minimum. Perceptron learning using backprop is an automatic way of setting the weights and thresholds. Training the perceptron to actual data allows the inference of the probability distribution of the data. Once the probability distribution is inferred, subsequent presentation of unseen inputs allows the perceptron to generalize and estimate an answer based upon inferred probability distribution.

2.7 Integration of Adaptive Optics with Artificial Neural Networks

This section answers research question 6 by describing how the previous implementations of artificial neural networks operated, the results of these implementations, and the limitations of these implementations. The earliest forms of the adaptive optical system control were realized using networks of operational amplifiers and resistors which achieved the desired solution during steady

state conditions (14:669). These implementations were a rudimentary form of the Hopfield artificial neural network (31:20). Recently, Lockheed presented further applications of artificial neural networks to the adaptive optical imaging system.

2.7.1 Lockheed Programmable Analog Neural Network. An analog neural network breadboard consisting of 256 neurons and 2048 programmable synaptic weights has been constructed and tested (7:260-268) (8:222-229). This analog neural network implements Hudgin's recursive algorithm for optimal wavefront estimation using an array of noisy wavefront slope measurements. The neural network finds the optimal control matrix to command the deformable mirror actuators. The actuators deform the mirror surface to compensate for the phase aberrated wavefront. The current implementations are an extension of work previously accomplished by Smithson in 1987 (31). In 1987, Smithson began his studies into the integration of artificial neural network with adaptive optics. Smithson noticed that the analog feedback network implementation of the adaptive optics control algorithm developed by Hudgin was a special case of the Hopfield neural network. The implementation of the Hudgin algorithm used fixed weights on the feedback actuator voltages characteristic of the Hopfield neural network applied to continuous work functions.

Smithsons' implementation was followed by work accomplished by Fisher and others. Fisher used Hudgin's algorithm to integrate artificial neural networks with an adaptive optics system. Fishers contributions was a demonstration of the ability for open and closed loop control (7:260-268) (8:222-229) of a segmented deformable mirror array.

The open loop control consisted of training the artificial neural networks using actual actuator command signals as the training vector with no feedback between actuator commands. Fisher's open loop control implementation is equivalent to a single-layer, feed-forward artificial neural network. Fisher's training used a perceptron type training algorithm to determine the best weights for the input slope sensor values. The open loop study was then extended to allow actuator command feedback, in this case both the input and feedback weights were trained using an algorithm similar to the backpropagation training algorithm. The closed loop control used actuator command feedback

and adapted the system to optical component drift or settling due to thermal effects. For closed loop control, the loop was opened at 3 second intervals and system parameters were sampled. Based upon the sampled parameters the input and output feedback artificial neural networks weights were adjusted.

The basis for the previous applications of artificial neural networks to the adaptive optical imaging system all relied upon the recursive algorithm developed in 1977 by Richard H. Hudgin (19:375-378). This algorithm uses a "large array" assumption. The algorithm also implicitly assumes that there is no significant interaction between the mirror actuators. These assumptions result in a convolutional form for a "perfect reconstructor" that results in a recursive algorithm. This recursive algorithm showed that an estimate of the actual wavefront phase is a linear combination of the four "nearest neighbor actuator" phase and phase difference values. The shortfall of this assumption is that it does not allow for edge effects encountered at the mirror periphery. Due to the noninteraction between deformable mirror commands, this algorithm is well suited to describe the operation of a segmented mirror, but is deficient when attempting to explain the operation of a continuous mirror.

2.8 Summary

The current adaptive optic imaging systems may not achieve the high data processing throughput requirements necessitated by future military image resolution requirements. The high degree of parallelism of neural networks has resulted in significant increases in the throughput, and may offer the solution to the computational frame time limitations encountered by adaptive optic imaging systems.

To extend the previous implementations of artificial neural network reconstructors, this thesis develops a two alternate techniques that allow full control of all the actuators on a continuous facesheet deformable mirror (research question 7). A recursive technique is developed that is well suited for a Hopfield artificial neural network. A feedforward technique is developed for a single layer perceptron artificial neural network architecture. Performance of the artificial neural network reconstructors is analyzed in response to selected parametric variations. The parameters that are

varied are the light level and atmospheric coherence diameter. The use of artificial neural network training is used to examine the performance enhancements achievable using artificial neural networks.

III. Theory

3.1 Introduction

The previous chapter provided background knowledge of the problem (atmospheric turbulence), the system (adaptive optical imaging system), and artificial neural networks. This chapter uses the background knowledge to develop artificial neural network reconstructors. This chapter outlines the approach taken to develop artificial neural network reconstructors. The basic approach consisted of the following research steps;

(1) Define the fundamental adaptive optical system models.

(2) Develop the comparison baseline (least squares reconstructor) model.

(3) Develop the reconstructor models that are implemented using artificial neural networks.

(4) Validate the reconstructor models by mapping the models into conventional artificial neural network state spaces.

This chapter is developed by sequentially addressing each research step. Section 3.2 satisfies research steps 1 and 2. Section 3.2 serves three primary purposes. The first purpose is to present the fundamental adaptive optical imaging system models. The second purpose is to derive the comparison baseline model (least squares reconstructor). The final purpose, and perhaps the most important is to present the primary design equation (Eq 3.9) used to develop the artificial neural network reconstructors. Research step 3 is addressed in section 3.3. Section 3.3 develops the recurrent and feedforward reconstructor models. Research step 4 is satisfied in section 3.4. Section 3.4 presents the conventional state space models for the artificial neural networks. The models developed in section 3.3 are then mapped into the artificial neural network state spaces.

3.2 The Adaptive Optics System Models

3.2.1 General System Development. This section satisfies research steps 1 and 2. First, the fundamental adaptive optical imaging system models are developed. Next, the development of the comparison baseline (least squares) reconstructor is presented. This section also presents the primary

design equation (Eq 3.9) for the artificial neural network reconstructors.

The adaptive optical system analysis closely follows the work done by Wallner (35:1771-1776), Welsh (38:1913-1923), and Roggemann (28:1-20). The analysis begins with the definition of the zero mean or piston removed wavefront phase $\phi(\underline{x},t)$, where $\underline{x}$ represents a two dimensional position vector within the system pupil and t is time. The piston removed phase, $\phi(\underline{x},t)$, is related to the phase of the incident wave front, $\psi(\underline{x},t)$, by

$$\phi(\underline{x},t) = \psi(\underline{x},t) - \int_{-\infty}^{+\infty} W_A(\underline{x})\, \psi(\underline{x},t)\, d^2\underline{x} \tag{3.1}$$

where $W_A(\underline{x})$ is the aperture weighting function. The aperture weighting function is assumed to satisfy the following relationship

$$\int_{-\infty}^{+\infty} W_A(\underline{x})\, d^2\underline{x} = 1 \tag{3.2}$$

This thesis assumes a circular aperture. A circular aperture results in an aperture weighting function of the form

$$\begin{aligned} W_A(\underline{x}) &= \frac{1}{\pi\left(\frac{D_o}{2}\right)^2} && |\underline{x}| \le \frac{D_o}{2} \\ &= 0 && |\underline{x}| > \frac{D_o}{2} \end{aligned} \tag{3.3}$$

where D_o is the aperture diameter.

The piston and tilt removed phase, $\gamma(\underline{x},t)$, is represented by the difference between the piston removed phase and the tilt component

$$\gamma(\underline{x},t) = \phi(\underline{x},t) - \underline{a}(t) \cdot \underline{x} \tag{3.4}$$

where the wavefront phase tilt component, $\underline{a}(t)$ is given by (28:6)

$$a(t) = \frac{16}{D_O} \int_{-\infty}^{+\infty} \underline{x}\, W_A(\underline{x})\, \phi(\underline{x},t)\, d^2\underline{x} \tag{3.5}$$

The analysis and simulation assumes that the piston and tilt component of the phase are removed.

A wavefront sensor is used to sense the wavefront phase $\gamma(\underline{x},t)$. The output of the wavefront sensor is modeled by (38:1914)

$$s_n(t) = \int_{-\infty}^{+\infty} W_n(\underline{x})\, [\nabla\gamma(\underline{x},t) \cdot d(\underline{x})]\, d^2\underline{x} + \alpha_n(t) \tag{3.6}$$

where $s_n(t)$ is the output of the n^{th} slope sensor at time t, $W_n(\underline{x})$ is the subaperture weighting function of the n^{th} slope sensor, $\nabla\gamma(\underline{x},t)$ is the spatial gradient of the piston and tilt removed phase, d is the unit vector in the direction of slope sensor sensitivity, and $\alpha_n(t)$ is the slope measurement noise for the n^{th} wavefront slope sensor.

The measurement noise, $\alpha_n(t)$ is due to shot noise effects in the wavefront sensor and is assumed to be uncorrelated with the wavefront phase. The noise is modeled as a zero mean Gaussian random variable. The standard deviation of the slope measurement noise, $\sigma_\alpha(t)$ is defined by (38:1918-1920)

$$\begin{aligned} \sigma_\alpha &= \frac{0.86\,\pi\,\eta}{\sqrt{N_p}\, r_o} && D_o \geq r_o \\ &= \frac{0.74\,\pi\,\eta}{\sqrt{N_p}\, D_o} && D_o < r_o \end{aligned} \tag{3.7}$$

where η is a correction factor which accounts for the finite sized detectors and N_p is the total

subaperture photon count. The value of η is assumed to be 1.5. This value corresponds to typical values found for charge coupled device technology (38:1919).

Integrating the first term in Eq 3.6 by parts, yields

$$s_n(t) = \int_{-\infty}^{+\infty} - [\nabla W_n(\underline{x}) \cdot d^n] \gamma(\underline{x},t)\, d^2\underline{x} + \alpha_n(t) \tag{3.8}$$

For notational simplicity, the bracketed term is represented by $W_n^s(\underline{x})$, where the superscript s represents the gradient of the weighting function in the direction of sensitivity of the n^{th} slope sensor.

The reconstruction matrix M_{jn} maps the measured slopes into a deformable mirror actuator command. This matrix is typically found by using a least squares or minimum variance approach (28:1)(24). This matrix is the essential element of the artificial neural network reconstructor. This matrix will determine the interconnection weights of both realizations of the artificial neural network reconstructors. The actuator command, $c_j(t)$ is related to the slope sensor measurements by

$$c_j(t) = \sum_n M_{jn}\, s_n(t) \tag{3.9}$$

Equation 3.9 is the fundamental design equation used to develop the artificial neural network reconstructors. Equation 3.9 determines how the measured wavefront slopes are used to produce the deformable mirror actuator commands.

The instantaneous phase correction associated with the actuator command $c_j(t)$ is defined by

$$\gamma_j(\underline{x},t) = c_j(t)\, e_j(\underline{x}) \tag{3.10}$$

where $e_j(\underline{x})$ is the actuator influence function or actuator impulse response function to a unit actuator command. Here and in the simulation Gaussian actuator influence functions are used. The actuator influence at any point $\underline{x}$ in the aperture plane due to the j^{th} actor is given by

$$e_j(t) = e^{-\frac{(\underline{x} - \underline{x}_j)^2}{2\sigma_j^2}} \tag{3.11}$$

where $\underline{x}_j$ is the location of the j^{th} actuator center, and σ_j is the rms width of the actuator influence function.

The sum over all the actuators of Eq 3.10 defines the estimated phase, $\gamma'(\underline{x},t)$ of the incoming wavefront.

$$\begin{aligned} \gamma'(\underline{x},t) &= \sum_j c_j(t)\, e_j(\underline{x}) \\ &= \sum_j \sum_n M_{jn}\, s_n(t)\, e_j(\underline{x}) \end{aligned} \tag{3.12}$$

3.2.2 Least Squares Reconstructors. The reconstruction matrix (M_{jn}) allows the calculation of the deformable mirror actuator commands using the measured slopes (Eq 3.9). The implementation of the conventional and artificial neural network reconstructors hinges on the definition of the reconstruction (M_{jn}) matrix. A conventional technique that defines the reconstruction matrix M_{jn} is the least squares approach.

The least squares or closed loop reconstructor uses feedback control of the actuator positions to compensate for the aberrated wavefront. The phase estimate, $\gamma'(\underline{x},t)$ is used to calculate an estimated slope measurement using Eq 3.8. The difference or error between the measured slopes ($s_n(t)$) and the estimated slopes is defined as s_n':

$$s_n'(t) = s_n(t) - \int_{-\infty}^{+\infty} -W_n'(\underline{x}) \sum_{j=1}^{J} [c_j(t)\, e_j(\underline{x})]\, d^2\underline{x} \tag{3.13}$$

Since the actuator command is independent of $\underline{x}$, the slope error can be written as

$$s'_n(t) = s_n(t) - \sum_{j=1}^{J} c_j(t) \int_{-\infty}^{\infty} -W'_n(\underline{x})\, e_j(\underline{x})\, d^2\underline{x} \tag{3.14}$$

The integral can be represented by a (J x N) matrix, P_{jn}, that maps the actuator command vector into subaperture space. The difference between the measured and estimated slopes is now

$$s'_n(t) = s_n(t) - \sum_n c_j(t) P_{jn} \tag{3.15}$$

To calculate the least squares reconstructor control matrix, the mean squared difference between the measured and the estimated slopes is minimized with respect to the actuator command vector. Minimizing the expansion of $s_n^T s_n'$ results in the least squares reconstructor control matrix

$$M_{jn}^{LS} = (P_{jn}^T P_{jn})^{-1} P_{jn}^T \tag{3.16}$$

Equations 3.14 through 3.16 define the terms of the least squares reconstruction matrix. The least squares reconstruction matrix is a function of only the actuator influence function and the directional gradient of the subaperture weighting function.

3.3 Reconstructor Model Development

This section satisfies research step 3, development of reconstructor models. The models developed in this section are implemented as artificial neural network reconstructors. To accomplish the development, Eq (3.9) is manipulated to forms easily implemented using recurrent and feedforward artificial neural networks. This section begins with the development of the recurrent reconstructor model and concludes with the feedforward reconstructor model.

3.3.1 Recurrent Reconstructor Model Development. This section uses equation 3.9 to develop a recurrent reconstructor model. This recurrent algorithm will then be mapped into Hopfield neural

network state space. Only after the algorithm is mapped into Hopfield state space can the name Hopfield neural network reconstructor be accurately applied.

The algorithm developed in this section forms the basis of the Hopfield neural network used in this thesis research. Equation 3.9 is manipulated to a form easily implemented using an recurrent artificial neural network. The feedback control algorithm development begins with the fundamental equation for the command signal to the j^{th} deformable mirror actuator (Eq 3.9).

$$c_j(t) = \sum_n \left[\left(M^x_{j,n}\, s^x_n(t) \right) + \left(M^y_{j,n}\, s^y_n(t) \right) \right] \tag{3.17}$$

where the contributions to $c_j(t)$ by the x and y directed wavefront sensor measurements are explicitly separated into two terms. Based upon this actuator command signal vector, the difference between the j^{th} any k^{th} actuator commands is

$$\begin{aligned} c_j(t) - c_k(t) = \sum_n \left[\left(M^x_{jn} - M^x_{kn} \right) s^x_n(t) \right. \\ \left. + \left(M^y_{jn} - M^y_{kn} \right) s^y_n(t) \right] \end{aligned} \tag{3.18}$$

Defining

$$\begin{aligned} \Delta M^x_{j-k,n} &= M^x_{jn} - M^x_{kn} \\ \Delta M^y_{j-k,n} &= M^y_{jn} - M^y_{kn} \end{aligned} \tag{3.19}$$

the difference in Eq 3.18 is rewritten as

$$c_j(t) = c_k(t) + \sum_n \left[\Delta M^x_{j-k,n}\, s^x_n(t) + \Delta M^y_{j-k,n}\, s^y_n(t) \right] \tag{3.20}$$

Using Eq 3.20 twice, the j^{th} actuator command can be written in terms of the k^{th} and the k'^{th} actuators

$$c_j(t) = \frac{1}{2} c_{k,n}(t) + \frac{1}{2} c_{k',n}(t)$$

$$+ \frac{1}{2} \sum_n \left[\left(\Delta M^x_{j-k,n} + \Delta M^x_{j-k',n} \right) s^x_n(t) \right. \tag{3.21}$$

$$\left. + \left(\Delta M^y_{j-k,n} + \Delta M^y_{j-k',n} \right) s^y_n(t) \right]$$

Equation 3.21 is the recurrent reconstructor model. Equation 3.21 will be implemented using a Hopfield artificial neural network. The difference in actuator commands between all the deformable mirror actuators and their neighbors defines a set of difference equations. These difference equations are used to establish interdependency of all deformable mirror actuator commands. The number of neighboring actuator commands used (i.e. how many feedback interconnections), is determined by the Hopfield neural network feedback interconnection weight matrix stability criteria (sect. 2.6.3).

The feedback interconnection matrix is defined by connecting a single actuator with only two of its neighbors. The interconnection weight of 1/2 from the j^{th} to the k^{th} actuator is equal to the interconnection weight from the k^{th} to the j^{th} actuator. The self feedback weight for the j^{th} actuator is zero. Knowing that two nearest neighbor actuator feedback commands are needed, the actuator commands are now linked by "connecting the dots" in the deformable mirror plane (Fig. 3.1). A single continuous line defines the actuator feedback connections. By using a single continuous line, the dependence of each actuator command upon all of the other deformable mirror actuator commands is defined. This dependence is insured through a cascading effect of actuator command dependence. An example of this 'cascading' effect is shown for actuator 1 whose command depends upon actuators 2 and 6. Actuator 6 also depends on actuator 12's command, and so on until each actuator command is related to a weighted version of each other actuator in the array.

3.3.2 Feedforward Reconstructor Model Development The algorithm used for the feedforward reconstructor model is simply Eq (3.9). Subsequent mapping of this equation into the single layer perceptron artificial neural network state space demonstrates the ease of implementation of this type of feedforward reconstructor model. The significant advantage of using a single layer perceptron artificial neural network over conventional techniques is the high degree of parallelism. This high

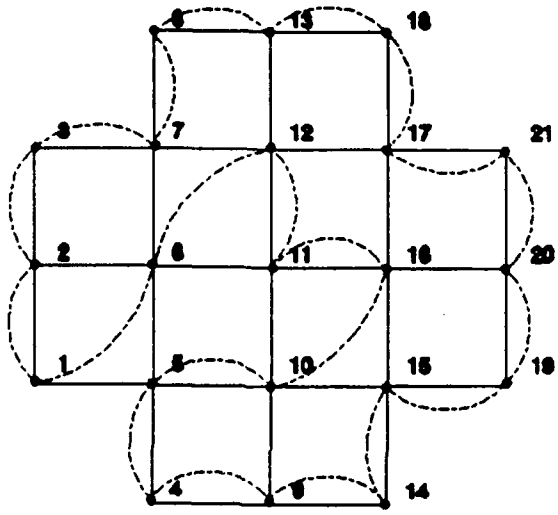


Figure 3.1 Actuator Feedback Connections

degree of parallelism should provide significant speed advantages. The added feature of artificial neural network training should enhance the performance of this technique over conventional reconstructors.

3.4 Reconstructor Model Mapping into Conventional Artificial Neural Network State Space

To satisfy research step 4, the reconstructor models must be mapped into conventional artificial neural network state spaces. This section answers the question, "Are these reconstructor models artificial neural networks ?" The mapping of the reconstructor models into neural network state spaces validates the models as being special cases of the Hopfield and single layer perceptron artificial neural network. Each reconstructor (feedforward, recurrent) mapping is developed by first

Table II Actuator Feedback Connections

j	k,k'	j	k,k'
1	2,6	12	6,11
2	1,3	13	8,18
3	2,7	14	9,15
4	5,9	15	14,19
5	4,10	16	10,11
6	1,12	17	18,21
7	3,8	18	13,17
8	7,13	19	15,20
9	4,14	20	19,21
10	5,16	21	17,20
11	12,16		

presenting the conventional neural network state spaces. The reconstructor models are then mapped into the conventional neural network state spaces. First, the feedforward model is mapped into single layer perceptron state space. Finally, the recurrent reconstructor model is mapped into Hopfield neural network state space.

3.4.1 Single Layer Perceptron. The general equation for a single layer perceptron artificial neural network can written as (27:52)

$$\bar{a} = W\bar{I} \tag{3.22}$$

where $\bar{a}$ is a vector of activations of the perceptrons, W is the weight matrix composed of each individual vector of perceptron weights, and $\bar{I}$ is the vector of inputs. When $\bar{a}$ represents the outputs of all the perceptrons, the single layer perceptron artificial neural network performs a mapping of the input vector into a set of output vectors.

3.4.2 The Feedforward Reconstructor Model and the Single Layer Perceptron. The feedforward reconstructor model is simply equation 3.9. Comparing equations 3.9 and 3.19 depicts a relatively simple mapping. For linear outputs, $\vec{a}$ corresponds to the vector of actuator commands, W corresponds to the reconstructor matrix M_{jn}, and $\vec{I}$ corresponds to the slope sensor inputs S_n. For sigmoidal activation functions, there must be an "inverse sigmoid" applied to each neuronal output. Sigmoid activation functions are used in this thesis to allow easier training of the single layer perceptron artificial neural network reconstructor.

3.4.3 The Hopfield Algorithm. This section defines the conventional Hopfield state space equations. For a continuous work function, Hopfield defines his neural networks equation of state as (18:3089:3090)

$$C_i \frac{du_i}{dt} = \sum_j T_{ij} V_j - \frac{u_i}{R_i} + I_i \tag{3.23}$$

where C_i is the input capacitance to a neuron, u_i is the input voltage to the i^{th} neuron, T_{ij} is the interconnection conductance (i.e. weight) from i^{th} neuron to j^{th} neuron, V_j is the output of j^{th} neuron, $R_i = 1/T_i$ is the input resistance to the i^{th} neuron, and I_i is the input current.

Upon convergence, or steady state conditions, Eq 3.23 reduces to

$$\frac{u_i}{R_i} = \sum_{j=1} T_{ij} V_j + I_i \tag{3.24}$$

3.4.4 The Recurrent Reconstructor Model and the Hopfield Algorithm. This section will map the recurrent reconstructor model into Hopfield state space. The steady state analysis (Eq 3.24) is used to define the Hopfield artificial neural network state space. Equation 3.21 shows that any actuator control signal is determined from a linear combination of two weighted neighboring actuator

feedback signals and a weighted sum of the slope sensor outputs. Upon manipulation of Eq. 3.21 the j^{th} actuator command is represented by

$$c_j(t+1) = \sum_k T_{jk}\, c_k(t) + \frac{1}{2} \sum_n \Delta M_{jn}\, I_n \tag{3.25}$$

where the T_{jk} is the multiplicative 1/2 term on the feedback actuator commands, the directional ΔM_{jn}'s are combined into a single ΔM_{jn}, and the directional slopes are combined and represented by I_n. The time dependence now reflects the discrete time synchrounous update nature of the model. A given actuator command is updated with the feedback information from the previous time. This results in the "t + 1" and "t" dependence of the actuator commands. The last term is a constant for any actuator until the slope sensor measurement are updated, therefore this term can be represented by I_j. The T_{ij} matrix is composed of weights equal to 1/2 for each neighbor of actuator j found in Table II. Therefore, at any time t+1, a general equation for the i^{th} actuator command due to feedback from the j^{th} neighbors and the weighted slope sensor outputs is written as

$$c_j(t+1) = f_h \left[\sum_j T_{ij}\, c_j(t) + I_j \right] \tag{3.26}$$

where f_h is the nonlinearity or activation function. This thesis uses a soft limiter with a unity gain in the linear region of operation. Here the assumption is made that the soft limiters do not saturate, therefore the right side of Eq 3.25 is just the argument of the limiter. Comparing Eq 3.25 to Eq 3.24 shows that they are virtually identical. For a linear activation function to the i^{th} neuron, the input is the same as the output, and the equations are identical except for the R_i term. The R_i is the input resistance for the i^{th} neuron. This thesis assumes that all of the neurons are identical, hence $R_1 = R_2 \ldots = R$. For this case, both sides of Eq 3.24 can be multiplied by R. The only effect of the multiplication by R is to scale the weight matrix and input vector to the Hopfield neural network, not affecting the convergence stability in any way. With these assumptions, the adaptive optical image reconstruction algorithm represents a special case of the Hopfield neural network.

3.5 Summary

This chapter has gone through the fundamental theory used in the development of this research effort. The reconstructor models conform to the Hopfield and the single layer perceptron artificial neural network models. Since the models conform to the conventional artificial neural network representations, the term "artificial neural network reconstructor" is now applicable.

IV. Analysis Results

4.1 Introduction

This chapter presents the results of the analysis and simulation. The task of analyzing and simulating the adaptive optical imaging system is an intricate process. To adequately address each specific part, the entire simulation process was divided into specific experiments. The resulting experiments are defined below.

Experiment #

(1) Establish system comparison baseline (least squares reconstructor) analysis results.

(2) Integrate, analyze, and test system performance using artificial neural network reconstructors. Compare performance of adaptive optical imaging system with artificial neural network reconstructors to baseline system.

(3) Examine the effects on adaptive optical imaging system performance by using artificial neural network reconstructors with sparsened slope measurement weights.

(4) Implement training algorithms and examine effects on overall system performance.

Section 4.2 discusses experiment 1. Section 4.2 presents the parameters used and the results obtained using the least squares reconstructor. Experiment 2 is discussed in section 4.3. Section 4.3 validates the performance of the artificial neural network reconstructors. The performance validation is obtained by comparing the artificial neural network reconstructor performance to the baseline established in experiment 1. Section 4.4 addresses experiment 3. Section 4.4 examines the effects of slope measurement weight sparsening on the artificial neural network reconstructors performance. Sparsening is accomplished by "zeroing out" two and four slope measurement weights to each neuron. The final experiment is examined in section 4.5. Section 4.5 examines the impact of neural network training on system performance.

4.2 Experiment #1 - Establishing a Comparison Baseline

The analysis and simulation began with the establishment of the comparison baseline. The comparison baseline system uses the least squares reconstructor linear estimation technique. The comparison baseline results are generated from an unmodified version of the simulation code provided by the Phillips Laboratory, Kirtland AFB NM.

4.2.1 System Parameters. The adaptive optical system parameters used for the least squares and artificial neural network reconstructors are shown in Table I. The variations in light level were chosen to encompass a wide range of operation. The range chosen is consistent with the light level variations examined by Welsh (38). The changes in atmospheric coherence diameter were based upon the many factors. The range of coherence diameter variations coincides with the range presented by Goodman (13). The actual values for the coherence diameters were chosen to examine the effects of partial coherence and small aperture diameter to r_o ratios.

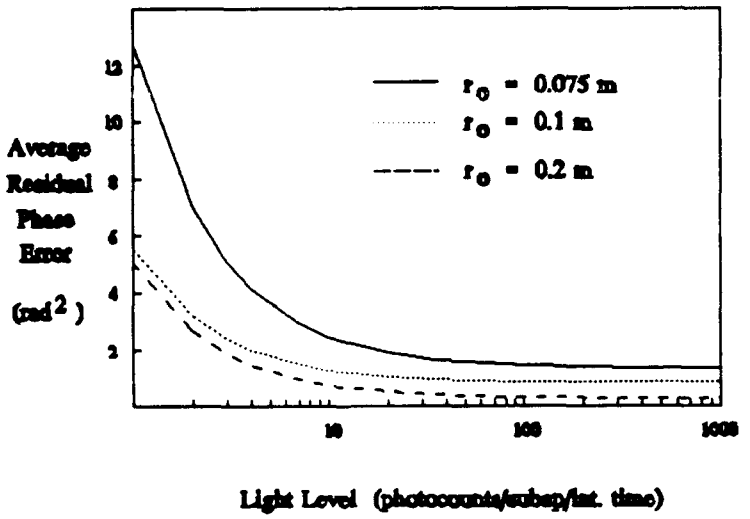


Figure 4.1 Baseline Adaptive Optical System Results - Least Squares Reconstructor

4.2.2 Baseline System Results. The results of the analysis of the baseline adaptive optical system (least squares reconstructor) is shown in Figure 4.1. The baseline or least squares reconstructor uses eq 3.9 to calculate the deformable mirror actuator commands. The response of the linear estimator (Fig 4.1) is consistent with the predicted system sensitivity. The establishment of this system baseline can now be contrasted to the performance obtained using artificial neural network reconstructors. The results for this comparison are presented in subsequent sections.

4.3 Experiment #2 - Artificial Neural Network Reconstructor Results

This section presents the analysis results for both artificial neural network reconstructors. The single layer perceptron results are presented first, followed by the analysis results for the Hopfield artificial neural network reconstructor.

4.3.1 Single Layer Perceptron Artificial Neural Network Reconstructor. The results of section 3.4.2 showed that the single layer perceptron artificial neural network reconstructor mapped into the exact same form as the least squares reconstructor. Therefore, the results shown in Figure 3.1 also represent the results obtained using a single layer perceptron artificial neural network reconstructor.

4.3.2 Hopfield Artificial Neural Network Reconstructor. The recurrent artificial neural network reconstructor developed in section 3.4.4 was integrated into the adaptive optical system simulation code. The Hopfield neural network is represented by a discrete time, synchronously updated system. The discrete time nature implies that the time intervals used to perform calculations and updates were discrete. The feedback values were synchronously updated to each Hopfield neuron. The Hopfield artificial neural network reconstructor was subjected to the exact same parametric variations as the baseline least squares reconstructor.

The results for the Hopfield neural network are grouped according to the atmospheric coherence diameter used. The first case used an atmospheric coherence diameter of 7.5 cm (Case 1). The general form of the reconstructor followed the baseline reconstructors form. The Hopfield

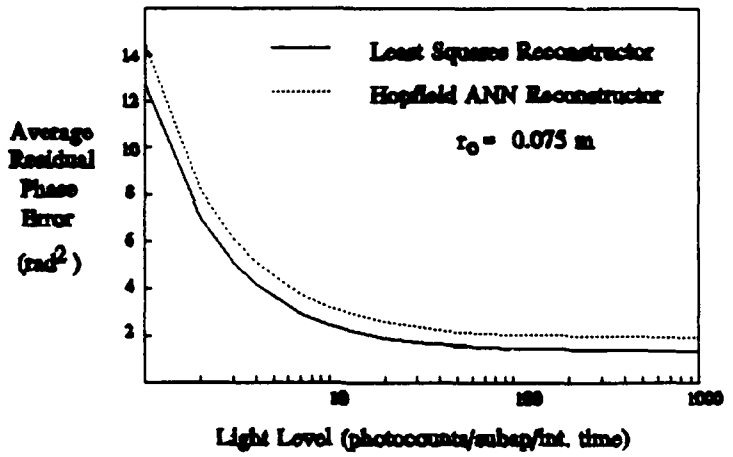


Figure 4.2 Experiment 2 - Case 1 Results

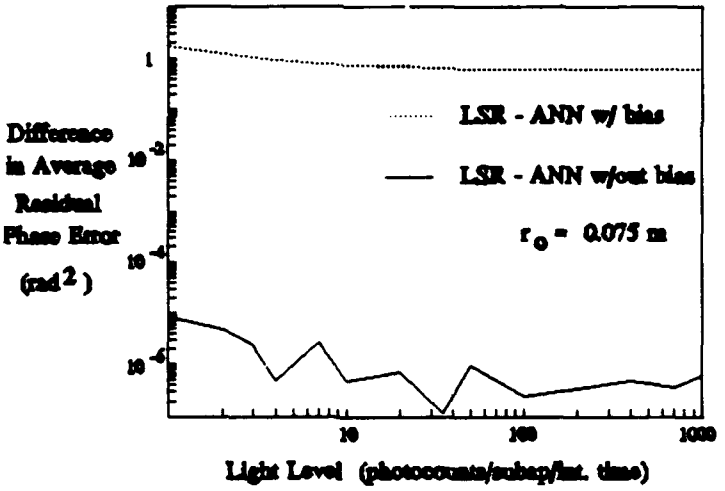


Figure 4.3 Experiment 2 - Case 1 Results, Effects of Bias Compensation

artificial neural network reconstructor resulted in a bias offset to each deformable mirror actuator command. The bias offset represents a piston or constant phase error to the reconstructed phase. The presence of a piston error to an adaptive optical imaging system is of no consequence. The effect of removing the piston error is shown in Figure 4.3. Figure 4.3 shows the difference between the least squares reconstructor and the Hopfield neural network reconstructor with and without the piston error. A constant piston error was removed from each actuator after the Hopfield neural network

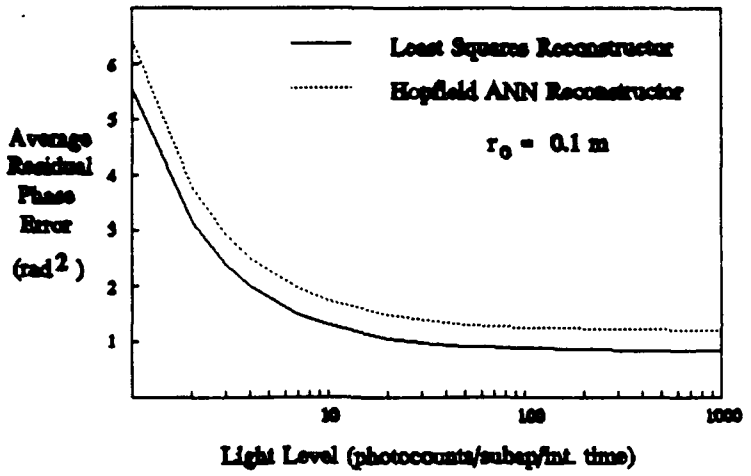


Figure 4.4 Experiment 2 - Case 2 Results

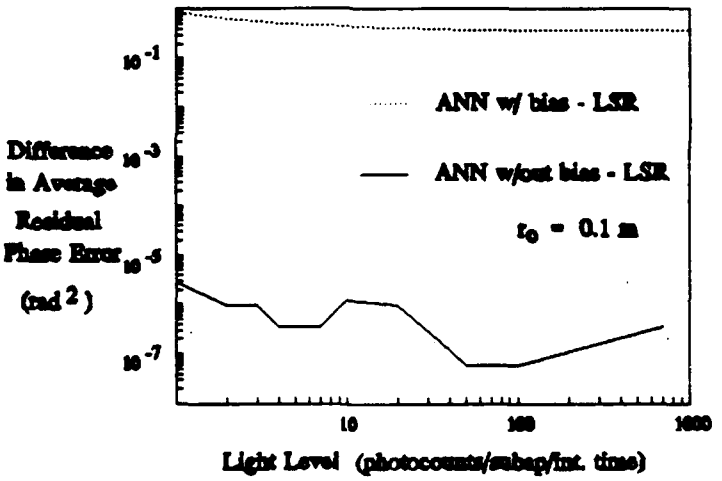


Figure 4.5 Experiment 2 - Case 2 Results, Effects of Bias Compensation

converged. Fig 4.3 shows that the average difference between the Hopfield artificial neural network reconstructor and the least squares reconstructor when the bias is removed is quite small. Therefore, the Hopfield artificial neural network reconstructor with bias removal obtains performance commensurate with the least squares and the single layer perceptron reconstructor. Figs. 4.4 through 4.7 show similar results for different atmospheric coherence diameters. Case 2 coincides with an r_o of 10 cm and Case 3 coincides with an r_o of 20 cm.

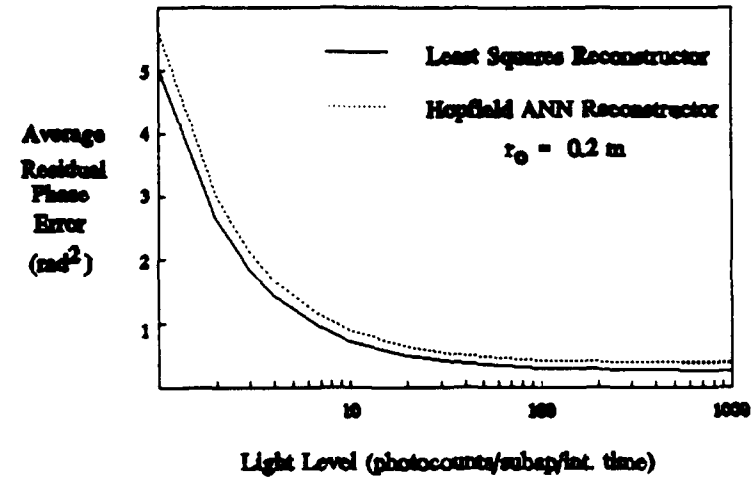


Figure 4.6 Experiment 2 - Case 3 Results

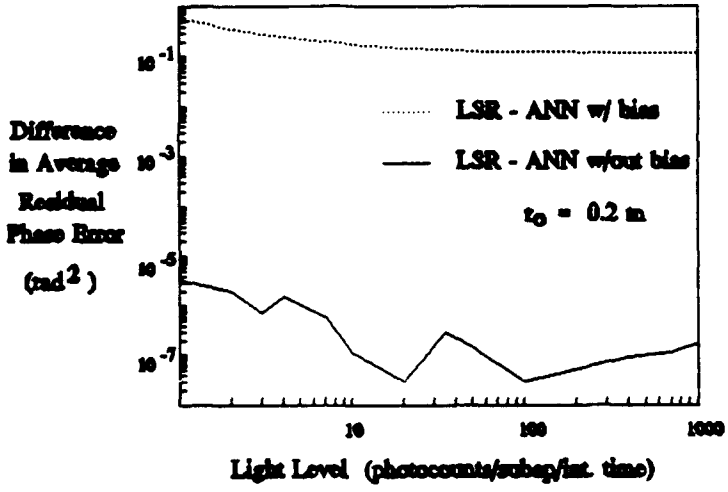


Figure 4.7 Experiment 2 - Case 3 Results, Effects of Bias Compensation

Although a constant piston error is insignificant in an adaptive optical imaging system, the origin of the error must be analyzed. The origin of the piston error was studied by examining all of the actuator values upon network convergence. This analysis found that each actuator command was offset from the least squares reconstructor actuator command by a constant value. The origin of the bias offset can have two possible explanations. First, the energy function of the Hopfield artificial neural network reconstructor is examined in Appendix A. The results of analysis show that the size

neural network reconstructor is examined in Appendix A. The results of analysis show that the size of the steps taken toward a energy minimum becomes increasingly smaller as the minimum is approached. This effect results in a small finite error between the actual actuator command and the actuator command the network converges to. Subsequent sections (Sect. 4.3.2) show that a change in system inputs to a Hopfield neural network may result in large changes in the system output. The effect of the convergence properties coupled with the recurrent nature may result in deviations in the system output from desired outputs. The second cause may be the lack of a constant bias compensation term to each Hopfield neuron. The effects of a constant bias compensation term have been explored in response to the presence of multiple minimas (30). This technique of minima compensation is known as "simulated annealing.". The technique of simulated annealing has been shown to produce a single global minimum. Simulated annealing is implemented in the Hopfield artificial neural network by introducing a bias current to each Hopfield neuron. A similar technique could be used to reduce the bias in the networks examined for this thesis.

4.4 Experiment #3 - Examination of the Effects of Sparsening

This section examines the effects of reducing the number of slope measurements available to each artificial neural network neuron or "sparsening." Each neuron must now calculate an actuator command using less information. The intent of sparsening is to reduce the number of interconnections required. This reduction in interconnections should equate to an overall reduction is system complexity and computational frame time.

The method used to determine which slope inputs to neglect was made by examining the weights on the slope measurements to the Hopfield artificial neural network. This method neglected specific slope sensor measurements by zeroing out the least significant weights to each neuron. Table III shows which slope sensor subaperture measurements were neglected. Table III shows which 2 ($S_n(2)$) and 4 ($S_n(4)$) slope sensor subaperture measurements that were neglected for each actuator. For example, for actuator 1, the least significant weights were those connecting slope measurements 4 and 8 to the 1st neuron. The four least significant weights were the weights connecting the 2, 3, 4,

Table III Sparsened Slope Measurement Data

Actuator #	S_n (2)	S_n (4)	Actuator #	S_n (2)	S_n (4)
1	4,18	2,3,4,18	12	18,19	5,4,19,20
2	4,18	4,12,16,18	13	15,18	8,12,15,18
3	12,18	4,12,16,18	14	2,7	1,2,6,7
4	1,3	1,3,4,13	15	2,6	1,2,5,6
5	4,18	2,3,4,18	16	15,19	15,16,18,19
6	3,18	3,4,8,18	17	5,15	5,10,15,18
7	12,18	8,12,16,18	18	15,18	5,10,15,18
8	12,18	8,12,15,18	19	5,11	1,5,6,11
9	2,11	2,6,7,11	20	5,7	5,6,7,11
10	10,18	2,5,10,18	21	5,15	5,10,11,15
11	18,19	15,18,19,20			

and 18th slope measurements to the 1st neuron. The same slope measurements that were neglected for the Hopfield artificial neural network reconstructor were also neglected for the single layer perceptron reconstructor.

4.4.1 Single Layer Perceptron. The effects of sparsening is examined for the single layer perceptron. The weights were set to zero for the specific sparsening case. The zero weights corresponds to zeroing out the appropriate M_{jn} matrix coefficients. The results of the sparsening of the single layer perceptron network is shown in Figure 4.8. Figure 4.8 shows that the performance characteristics of the sparsened single layer perceptron artificial neural network are degraded.

4.4.2 Hopfield The results for the sparsened Hopfield artificial neural network is shown in Figure 4.9. For the Hopfield artificial neural network, the sparsening did not correspond to zero out the M_{jn} coefficients. The determination of the input weights is attributed to the lack of one to one correspondence between input weights to the Hopfield artificial neural networks and the M_{jn} matrix

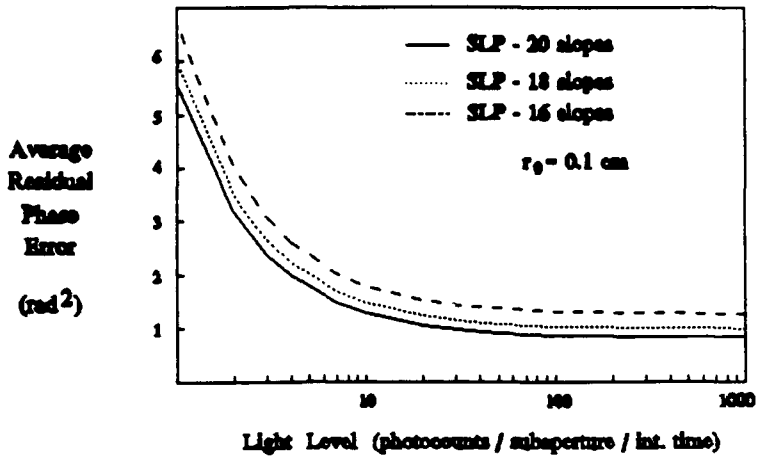


Figure 4.8 Experiment 3 - Sparsened single layer perceptron results

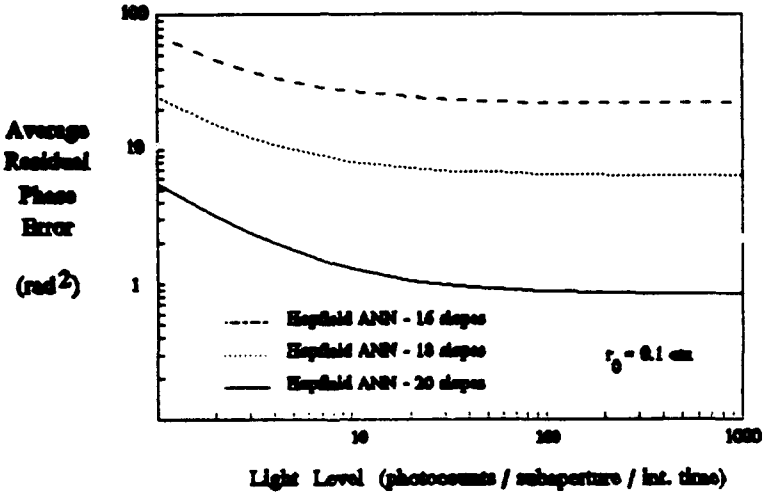


Figure 4.9 Experiment 3 - Sparsened Hopfield artificial neural network results

coefficients. The sparsening of the Hopfield artificial neural network reconstructor zeroes out the $_{\Delta}M_{jn}$ coefficients (Eq 3.21). These zeroed out coefficients are combinations of the differences in the M_{jn} matrix and a shifted version of the M_{jn} matrix. The actual effects of zeroing out the M_{jn} coefficients (single layer perceptron sparsening) and producing the $_{\Delta}M_{jn}$ matrix was examined. For example, in some cases zeroing out 4 M_{jn} coefficients and then producing the $_{\Delta}M_{jn}$ matrix resulted in the zeroing out of only 2 $_{\Delta}M_{jn}$ coefficients.

4.5 Experiment #4 - Training Results

The input weights to the single layer perceptron and the Hopfield artificial neural network reconstructors were trained. The goal of training was to determine if artificial neural network training could compensate for the lack of information presented by sparsening. The algorithms used to train the artificial neural network reconstructors are found in Appendices B and C.

4.5.1 Single Layer Perceptron Artificial Neural Network Reconstructor. The training results of the single layer perceptron artificial neural network reconstructor are depicted in Figure 4.10. The

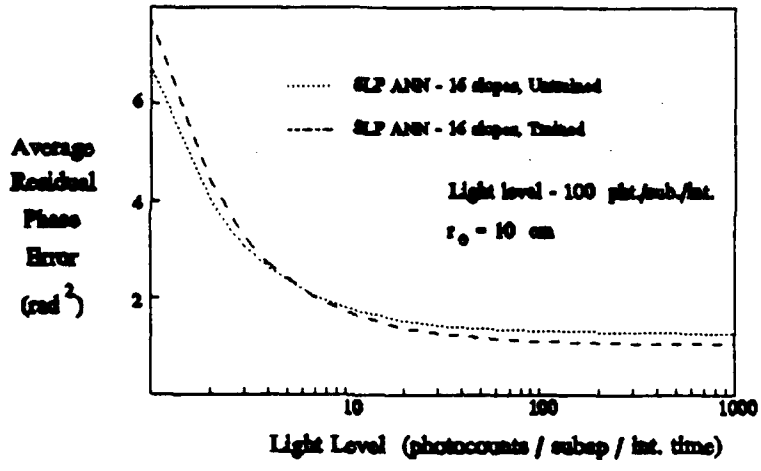


Figure 4.10 Experiment 4 - Single Layer Perceptron Training Results

single layer perceptron was trained using a training and test set composed of 400 training pairs per epoch. The light level was 100 photocounts/subaperture/integration time, and the atmospheric coherence diameter was 10 cm. Figure 4.10 shows that the performance of the sparsened artificial neural network reconstructor was improved by training. At light level above approximately 7 photocounts/subaperture/integration time, the trained artificial neural network provided better compensation for wavefront aberrations. The lack of performance of the trained artificial neural network reconstructor at low light levels is explained by the ability of the net to learn the probability

distribution of the noise. At high light levels, the noise variance is smaller than the noise variance at low light levels. Hence, training at high light levels would teach the network the noise variance properties using small variances. You could not expect the network to provide accurate estimates when the noise variance gets markedly larger than from what the network learned.

4.6 Summary

This chapter presented the analysis and simulation results. Section 4.3 showed that the artificial neural network reconstructors could achieve performance commensurate with conventional linear estimation techniques. In an effort to reduce the number of hardware interconnections, the effects of sparsening were examined in section 4.4. The result of section 4.4 was that the sparsened recurrent artificial neural network reconstructor experienced a larger performance degradation. The goal of section 4.5 was to find out if artificial neural network training principles could be applied to compensate for input sparsening. Section 4.5 showed that a reduction in hardware interconnections could be coupled with backpropagation training principles. The result was performance enhancements not currently achievable using conventional linear estimation techniques.

V. Conclusions and Recommendations

5.1. Introduction

This thesis researched the implementation of baseline artificial neural network reconstructors with an adaptive optical imaging system. This research examined the performance of the single layer perceptron and the Hopfield artificial neural network reconstructors. The reconstructors were integrated with an adaptive optical imaging system that was composed of a wavefront sensor with 20 slope sensor subapertures and a 21 individually controlled actuator deformable mirror.

The entire process was written in Fortran 77 and simulated on a VAX VMS system. The main simulation code was provided by the Phillips Laboratory, Kirtland AFB NM. The software written for this thesis consisted of the code to calculate the Hopfield input weights, the artificial neural network reconstructor subroutines, and the artificial neural network training algorithms.

This chapter presents a brief summary of what was accomplished by this research effort. This summary of accomplishments is followed by a summary of recommendations for future research into this area.

5.2. Conclusions

1. The artificial neural network reconstructor is a feasible alternative to conventional linear estimation techniques. The artificial neural network is an extremely fast, densely packed, highly parallel distributed processing structure composed of many simple processing elements. These features alone merit consideration of the artificial neural network as an alternative to conventional linear estimators. This research also demonstrated that there is virtually no performance difference between artificial neural networks and the conventional least squares reconstructor.

2. The added feature of artificial neural network training gives the artificial neural network the ability to provide performance enhancements not currently achievable using conventional

linear estimators. This performance enhancement was demonstrated by the trained single layer perceptron using only 16 of 20 slope sensor inputs.

3. The recurrent artificial neural network can provide performance commensurate with conventional linear estimators and single layer perceptrons. The feedback nature of recurrent neural networks did not result in the system robustness originally envisioned. Instead, the recurrent nature may result in inherent instabilities and a reduction in the systems ability to compensate for the lack of essential information.

5.3 Recommendations

This research represented an initial effort at AFIT to integrate two diverse technologies; adaptive optics and artificial neural networks. A general recommendation is that the entire simulation should be moved to a workstation environment. Throughout this effort, the mainframe computer reliability was quite unpredictable. The only pitfall is that a workstation environment would require rewrite of the IMSL subroutines used by the simulation. A residual effect of this groundbreaking effort was many simplifications were made to complete the research in the allotted time. Therefore, the recommendations for future research are many and grouped into two basic areas; adaptive optical imaging systems, and artificial neural networks.

5.3.1 Adaptive Optical Imaging System. This research was based upon a nominal amount of wavefront slope sensors and deformable mirror actuators. A logical extension is to increase the amount of slope sensors and deformable mirror actuators. By increasing the both the amount of slope sensor subapertures and deformable mirror actuators, the effects of reducing the number of slope sensor inputs or "sparsening" can be further researched. Coupled with the increase in the slope sensor subapertures and deformable mirror actuators should be an increase in the mirror diameter.

Further research should examine different techniques to reduce the amount of slope information used by each neuron. This research used the weights of each slope sensor measurement to determine the importance of each slope measurement to a particular neuron. Spatial location should be explored as means of reducing the amount of slope measurements input to each neuron.

The mirror surface may not be flat when actuator voltages are set to zero. Additional mirror data is needed to model these effects. A bias may be added to selected deformable mirror actuator commands to compensate for this effect. The effect of distinct actuator influence functions can be modeled.

The temporal aspects of the system also need to be researched. An instantaneous correction was assumed for this thesis. The effects of finite and random time delays between actuator command and movement need to be addressed. The temporal aspects of the wavefront sensing process need to be analyzed. The Hartmann sensor can undergo slow transients when detecting time invariant phase distortions. These slow transients are attributed to the sensor "warming up". These effects can be modeled as a random bias for each slope channel.

A hardware system composed of a continuous facesheet deformable mirror and a Hartmann wavefront sensor is available for use to provide verification of the software simulation results. Operational testing and evaluation of this system should provide additional data that can be integrated with the software simulation.

5.3.2 Artificial Neural Networks. The baseline artificial neural networks studied in represented two representative networks for two classes; recurrent and feedforward. Additional research into additional types of artificial neural networks is highly recommended. Many fundamental questions still remain unanswered. For instance, historically there has been two types of linear reconstructors; minimum variance and least squares. Can an artificial neural network somehow be trained to provide performance commensurate with the

minimum variance but have the robustness of the least squares reconstructor ? Can a neural network compensate for idiosyncrasies of the system components through training. Research into multilayer perceptron type networks is essential. Recently, a recurrent backpropagation network was demonstrated at AFIT with great success. The possibility of using this type of network to perform wavefront reconstruction is encouraged. This thesis outlined the fundamental implementation of the artificial neural network into the adaptive optical imaging system. The fundamental goal of future research should be geared to find the best candidate neural networks for implementation into the adaptive optical imaging system environment.

The artificial neural networks used in this thesis effort did not use a bias term into each neuronal element. This bias term should be added to each neuron of both the recurrent and the Hopfield artificial neural networks. The effects of this bias on the weight correlation to known reconstructor matrix coefficients should be examined. For classifiers, the bias allows partitioning of the feature space by offsetting the hyperplane locations. For function estimation, the effect of this bias term must be analyzed in depth. The effect of the bias on training should also be addressed. Training the network without bias is analogous to "tieing one hand behind the networks back". The bias adds an additional degree of freedom to the training algorithms. This added degree of freedom and its effect on training time and learning rates should be analyzed. Analysis into the use of backpropagation with momentum and conjugate gradient training techniques is encouraged.

Software simulation is meant to be a predictor of actual hardware implementations. A hardware implementation is now possible and encouraged here at AFIT. Neural network integrated circuits are currently available, coupled with the deformable mirror and Hartmann wavefront sensor available at AFIT, increased research into the spatial and temporal aspects is essential. Concurrence of hardware analysis and testing with software simulation is encouraged.

Appendix A: Hopfield Artificial Neural Network Energy Function Analysis

A.1 Introduction

The convergence of this type of neural network is essential to accurate wavefront phase estimates. The energy function for a Hopfield neural network describes the network convergence properties. Close examination of the gradient of the energy function defines the convergence properties for this type of neural network. The energy function of the Hopfield neural network with a continuous work function has unique properties. The energy function of the Hopfield neural network with a continuous, linear activation function is described by Eq A.1 (18:3089-3090). Using the same notation as section 3.4.3, for a unity gain, the inverse I/O relationship reduces to simply V_i, which gives us

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij} V_i V_j + \sum_{i=1}^{N} \frac{V_i^2}{R_i} + \sum_{i=1}^{N} I_i V_i \tag{A.1}$$

Examination of the gradient of the energy function with respect to the input voltage V_i, yields

$$\frac{dE}{dV_i} = -\frac{1}{2}\sum_{j=1}^{N} T_{ij} V_j + \frac{2V_i}{R_i} + I_i \tag{A.2}$$

Analyzing this definition of the change in the energy with respect to a change in the neuronal input yields interesting results. For identical neurons with a constant input resistance R, the right hand side of the gradient equation results in the three cases shown in Eq (A.3). The first case results in a negative gradient. The second case occurs at convergence for the energy surface or a minimum. The third case results in a positive gradient. For each case, the magnitude of the gradient is equivalent to the difference between the left hand side and the right hand side of each case.

$$(1)\quad \frac{2}{R} V_i + I_i < \frac{1}{2} \sum_{j=1}^{n} T_{ij} V_j$$

$$(2)\quad \frac{2}{R} V_i + I_i = \frac{1}{2} \sum_{j=1}^{n} T_{ij} V_j \qquad \text{(A.3)}$$

$$(3)\quad \frac{2}{R} V_i + I_i > \frac{1}{2} \sum_{j=1}^{n} T_{ij} V_j$$

The form of the energy surface is found by applying the second derivative test to the energy gradient. The second derivative of the energy function yields $2/R_i$. Since the resistance is always positive, the energy function is concave upwards. The critical values for the second derivative test are found by solving for the energy gradient of zero , Case 2 . The second derivative test reveals that these critical values represent minimums on the energy surface. Figure 3.1 is a graphical representation of the

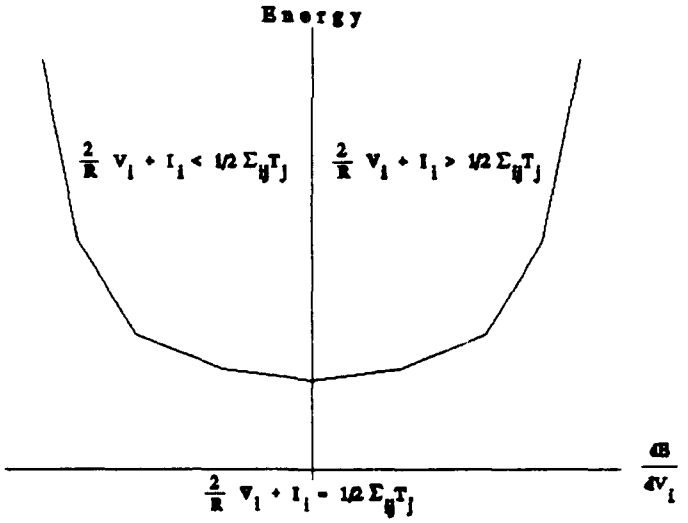


Figure A.1 Energy Surface Representation

energy surface outlined by the previous analysis.

The energy function provides a gradient descent down the energy surface toward a minimum. The magnitude of the gradient determines the size of the steps taken toward a minimum. Therefore, as the difference in the weighted output versus the weighted inputs decreases (Eq A.3 cases 1,3), the movement down the energy surface decreases accordingly. The Hopfield network step size is also

decreasing until a minima is reached. This step size reduction results in a slowing down of the movement down the energy surface as the minima is approached.

A useful contrast is made by examining the associative memory application of the Hopfield neural network. The step size of an associative memory is based upon the relationship of the energy gradient to a threshold value. The hard limiter used in the associative memory Hopfield neural network results in a discrete step size. The step size of the associative memory Hopfield neural network has values of 0,1,2. The discrete step size results in extremely fast convergence toward a minimum. The Hopfield network studied for this thesis has continuous outputs, this results in a larger variation in the step sizes taken toward an energy minimum. The variations in step sizes coupled with the decrease in step sizes as the minimum is approached has resulted in an increase in the number of iterations taken to converge. The problem with this type of minimum seeking algorithm is the presence of local minimums.

A.2 Uniqueness of Hopfield Network Solution. This type of neural network reconstructor usefulness is based upon the convergence to an accurate phase estimate. Zhao and Mendel (41:499-506) proven that a neural network with a specific form will converge to a unique solution defined by the initial conditions. A unique output is an output with a single or global minimum. What this infers, is that under specific conditions this class of neural network will calculate a unique answer for unique inputs. For the adaptive optical imaging system, this equates to a unique phase estimate for a unique set of system conditions.

The form of the network examined by Zhao and Mendel corresponds to a Hopfield neural network, with specific assumptions. The assumptions employed by Zhao and Mendel apply to the neural network studied for this thesis and are as follows

(1) Feedback interconnection matrix is symmetric

(2) No self feedback for individual neurons

(3) Initial condition for the neuronal outputs is zero

(4) Neuron outputs are bounded within a well defined region

(5) Linear gain within the output boundary region

(6) All neurons have the same I/O characteristics.

The lack of local minimums alleviates the concern that the neural network reconstructor may provides erroneous actuator commands.

Appendix B: Single Layer Perceptron Artificial Neural Network Training

B.1 Introduction

This appendix examines the training technique used for the single layer perceptron artificial neural network reconstructor. The artificial neural network training used is the backward error propagation technique or simply "backprop." This thesis used the instantaneous method of backpropagation training. The instantaneous method minimizes the error associated with a single training exemplar. Weight updates are made for each training exemplar presented to the artificial neural network.

B.2 Instantaneous Backpropagation Method

The instantaneous backprop method begins by defining the total error for the j^{th} neuron, ε_j as

$$\varepsilon_j = (d_j - c_j) \tag{B.1}$$

where d_j is the desired output and c_j is the calculated output for the j^{th} neuron. The total energy (E_{tot}) is defined as the sum of the individual mean squared errors or

$$E_{tot} = \sum_j \frac{1}{2} (d_j - c_j)^2 \tag{B.2}$$

The weight updates are implemented using the first-order gradient steepest descent (27:103.) The weight update equation for the input interconnection weights (w_{ij}) is defined by

$$\begin{aligned} w_{ij}^{+} &= w_{ij}^{-} + \Delta w_{ij} \\ \Delta w_{ij} &= -\eta \frac{\partial E}{\partial w_{ij}} \end{aligned} \tag{B.3}$$

The new weight, w_{ij}^{+}, is written in terms of the old weight, w_{ij}^{-}, the energy, and a learning rate, η. The single layer perceptron network is depicted in figure 1.5. The task that remains is the calculation of the weight change, Δw_{ij}. Using equation B.2, the change in weight becomes

$$\Delta w_{ij} = - \eta \ (d_j - c_j) \ \frac{\partial c_j}{\partial w_{ij}} \tag{B.4}$$

The calculated output of the single layer perceptron, c_j is define by

$$c_j = f_h \ (\sum_i \ w_{ij} \ I_i + \theta_j)$$

where f_h is the activation function or nonlinearity, I_i is the i^{th} input feature to the j^{th} neuron, and θ_j is the bias term. For this thesis the activation function used is the sigmoid or "squashing" function. When the argument of the activation is represented by $\alpha(w_{ij})$, the calculated output using a sigmoidal activation function is now

$$c_j = \frac{1}{1 + e^{-\sum_i w_{ij} I_i + \theta_j}} \tag{B.6}$$

The partial derivative in equation B.4 is now

$$\frac{\partial c_j}{\partial w_{ij}} = c_j \ (1 - c_j) \ I_i \tag{B.7}$$

Using equations B.7 and B.3, the equation for the weight updates becomes

$$w_{ij}^{+} = w_{ij}^{-} + \eta \ (d_j - c_j) \ c_j \ (1 - c_j) \ I_i \tag{B.8}$$

To implement this algorithm, the following steps were used;

(1) Initialize weights and thresholds to small random numbers.

(2) Statistically normalize the input features

(3) Present training vector and desired output

(4) Calculate the actual output using B.5

(5) Update weight using B.8, where $0 \leq \eta \leq 1$

Appendix C: Hopfield Artificial Neural Network Training

C.1 Introduction

This thesis will examine if the conventional Hopfield neural network can be trained to provide performance enhancements under varying system and hardware conditions. The algorithm developed as part of this research effort begins with the conventional backpropagation algorithm.

The conventional Hopfield artificial neural network utilizes the feature of fixed interconnection weights to provide extremely fast convergence to a minimum on the energy surface. The recurrent backpropagation neural network is a special case of the Hopfield neural network. The main distinctions between the Hopfield artificial neural network and the recurrent backpropagation neural network are that the recurrent backpropagation neural network uses a conventional backpropagation training algorithm to adapt both feedback and feed forward (input) weights. The recurrent backpropagation neural network is not designed to iterate until convergence like a Hopfield neural network, but instead is developed as a predictor of future neuron output states. Due to this future state predictor application, the recurrent backpropagation neural network also assumes that there is some temporal relationship between input training exemplars. This thesis will develop a neural network that diverges from the conventional Hopfield and recurrent backpropagation neural networks. This hybrid network will have fixed feedback interconnection weights and will train the input weights. There are numerous reasons for this methodology. First, the feedback interconnection weight matrix is well defined. Altering of the feedback interconnection weight matrix may in fact jeopardize the stability criterion fundamental to this type of recurrent neural network. Another reason is that due to the sheer magnitude of the problem, the reduced weight training set will equate to a reduction in overall neural network training time. The general system to which the training algorithm is to be applied is shown in Fig (C.1).

There are fundamentally two ways to implement the backpropagation training algorithm, the batch and instantaneous methods. The batch backpropagation training algorithm calculates the ensemble training exemplar error, using the sum of the instantaneous errors ($\epsilon_k^p(t_e)$) over the training

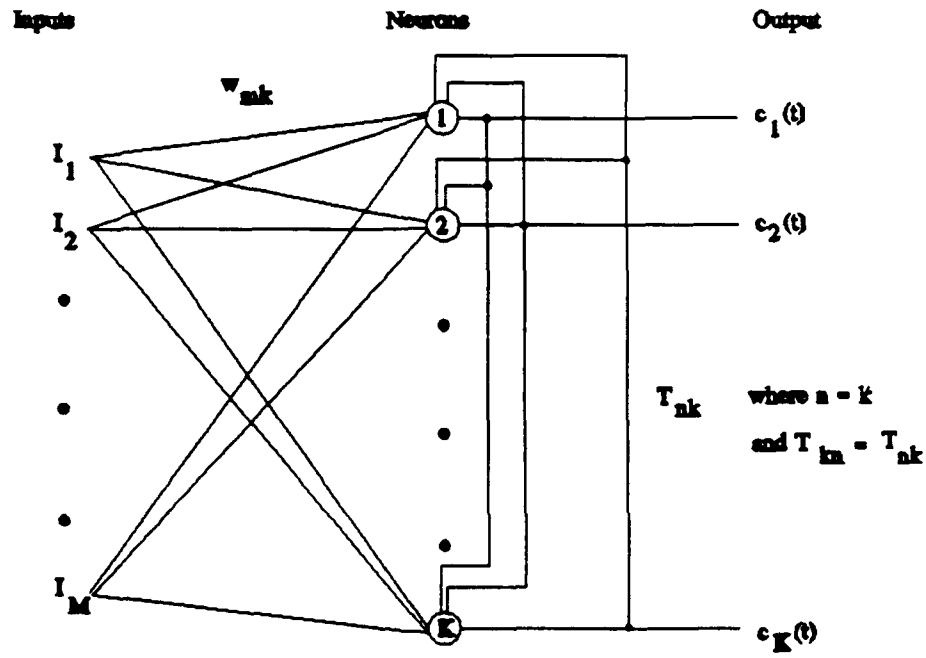


Figure C.1 Generic Recurrent Neural Network Architecture

set as the measure to minimize. The superscript p denotes an instantaneous value, and the subscript k denotes the k^{th} neuronal element. The t_e is the time when the network settles or converges to an answer. For this thesis effort, convergence stability is defined as 100 successive iterations where the new output is less than 0.1% different than the previous output. The batch backpropagation method defines the energy function for p training exemplars as

$$E_{batch} = \sum_{p=1}^{P} \sum_{k=1}^{K} \frac{1}{2} (\varepsilon_k^p(t_e))^2 \qquad \text{(C.1)}$$

The instantaneous error method seeks to minimize the error associated with a single training exemplar. In this case the energy function is defined as the sum of the squared error function.

$$E^p = \sum_{k=1}^{K} \frac{1}{2} (\varepsilon_k^p(t_e))^2 \qquad \text{(C.2)}$$

In both cases, the weight update algorithm uses a gradient descent down the error surface to converge toward a minimum.

C.2 Ensemble Training Technique

For the purposes of this thesis, the recurrent backpropagation learning algorithm will be employed, with a slight twist. The weight updates will not occur at every Hopfield iteration, but will be done after the network converges to a solution for a given input training exemplar. This is done because, changing the weights during Hopfield iterations, interrupts the recursion process and may result in nonconvergence of the neural network. This differs from the conventional recurrent backpropagation network training algorithm, which trains feedback and input weights after each iteration. A recurrent backpropagation network adopts this type of training because what is important is the temporal relationship between iterations, rather than the output upon convergence.

The Hopfield neural network is concerned with the output value at the time of convergence. For this application the neural net must be allowed to converge to a solution, whether right or wrong. The difference between this solution and the desired solution defines the instantaneous error. The instantaneous error for that exemplar is made up of contributions of error for each Hopfield iteration. The energy function analysis must then be propagated back through each iteration and the average weight update will be used to determine the new weights. The weight update for each exemplar, will be made prior to presentation of another training exemplar. This backpropagation of weight updates is represented by a total weight update equal to the mean of the weight updates that would have been done for each iteration. This thesis assumes that the neural network operational characteristics can be modeled by a discrete time, synchronously updated system. If the process begins at time t_o and ends at time $t=t_e=t_o+Q$, the change in weights per training exemplar is defined by

$$\Delta w_{ij}^{p}(t_e) = \frac{1}{Q} \sum_{q=1}^{Q} \Delta w_{ij}^{q}(t_o+q) \qquad \text{(C.3)}$$

At this point, it is important to remember that the weight update occurs after the network has converged ($t=t_e$), and not at each iteration. This type of weight update scheme is represented by

$$w_{ij}^{p} = w_{ij}^{p-1} + \Delta w_{ij}^{p} \tag{C.4}$$

For each Hopfield iteration, this system does not update the weights, but must keep track what the weight updates would have been. Each iteration calculates the present change in weights by using information from previous weight updates. This can be termed a "short term system memory", since knowledge of all training iterations is needed to calculate a weight update per training exemplar.

By examining the process in a sequential time sense, when a new exemplar is presented, the weight update for the past training session represents the point of origin for the weights for the new training exemplar. Each training exemplar defines a new energy function in energy/weight space. The gradient descent of the new energy function begins at the location of the last weights. Therefore, the calculation of the new weight update for the present training exemplar is a gradient descent to a minimum in reference to the last minimum. This sequential process of updating the weights for many exemplars can now be termed the "systems long term memory", since the sequential referencing of the present to the past weight update takes into account the effects of the weight updates for all the past training exemplars. A large ensemble of training exemplars will be presented to the neural network. The presentation of a large set of exemplars is attempting to build a global energy surface for the random process. By utilizing the long term memory of the system, it is hoped that generalization about the adaptive optics system statistics (global energy function) will occur. This generalization should result in the formation of an optimal set of weights for the wavefronts random process that correspond to the location of the global energy minimum in weight space. To determine the ensemble training algorithm, the iterative qualities of that ensemble training algorithm must be defined.

C.2.1 Iterative Aspects of the Ensemble Training Algorithm. To analyze the time dependent aspects of the Hopfield neural network we must first define the temporal transition characteristics of

the Hopfield neural network. To accomplish this, an analogous network to Fig (C.1) is shown in Fig (C.2).

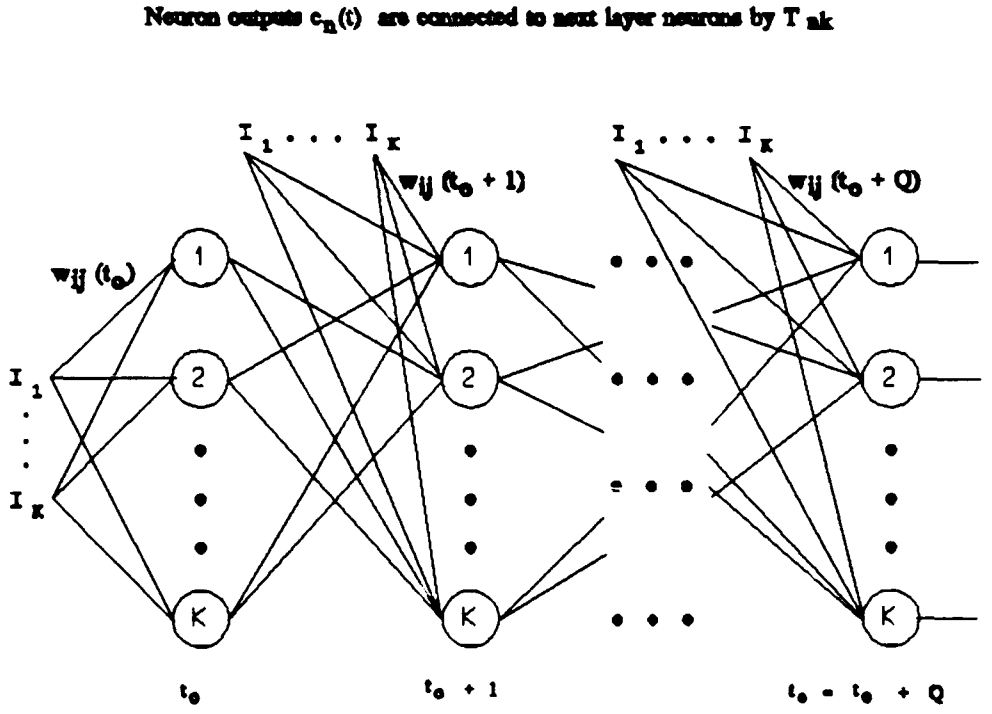


Figure C.2 Discrete Time Representation of the Hopfield Artificial Neural Network

Where the subscript Q represents the time at which convergence occurs. The development of the instantaneous backpropagation learning begins with the definition of the weight update equation at any time t, for the system defined in Fig (C.2).

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t) \qquad \text{(C.5)}$$

Using this notation, at a time t, w_{ij} is defined as the weight from the i^{th} input feature to the j^{th} neuron. Where $w_{ij}(t)$ represents the new weight, $w_{ij}(t-1)$ is the old weight, and the change in weights is denoted by $\Delta w_{ij}(t)$. The weight update at time t is defined as

$$\Delta w_{ij}(t) = -\eta \frac{\partial E^p}{\partial w_{ij}(t)} \qquad \text{(C.6)}$$

The negative of the gradient points in the direction of greatest decrease or descent on the energy surface. Thus the gradient vector provides direction along the surface of the energy function which searches for a minimum. The factor η is the gain or the learning rate of the training algorithm. Applying the chain rule to the gradient of the energy function with respect to the activation results in

$$\frac{\partial E^p}{\partial w_{ij}(t)} = \frac{\partial E^p}{\partial c_k(t)} \frac{\partial c_k(t)}{\partial w_{ij}(t)} \qquad \text{(C.7)}$$

For the adaptive optical imaging system implemented using a Hopfield neural network, the actuator command k^{th} neuron is defined as

$$c_k(t) = f_h\left[\sum_{n=1}^{K} T_{nk}\, c_n(t-1) + \sum_{m=1}^{M} w_{mk}(t)\, I_m^p\right] \qquad \text{(C.8)}$$

This simply states that the output of the k^{th} neuron at any time t is described by a nonlinearity whose argument is the sum of two terms. The first term is a linear combination of the weighted actuator outputs one time step earlier. The weights are defined by the feedback control matrix. For stability purposes, the main diagonal of the feedback interconnection matrix is set to zero and off diagonal terms are symmetric, i.e. $T_{nk} = T_{kn}$. The second term is the weighted sum of the m input features, this is the set of weights that are to be trained. Examining the partial derivative of the output of the k^{th} neuron with respect to the input weights results in

The derivative looks at the change in the activation of the actuator command at time t with respect to the weights at time t. Therefore the first term is independent of the change in weights at time t,

$$\frac{\partial c_k(t)}{\partial w_{ij}(t)} = \frac{\partial}{\partial w_{ij}(t)} f_h[\sum_{m=1}^{K} T_{mk}\, c_k(t-1) + \sum_{m=1}^{M} w_{mk}(t)\, I_m^p] \tag{C.9}$$

and equals zero. Examination of the second term results in a derivative that exists only when m=i, and k=j. The change in the activation of the actuator command with respect to the weights at time t can now be represented by

$$\frac{\partial c_k(t)}{\partial w_{ij}(t)} = \delta_{jk} I_i^p \tag{C.10}$$

Referring back to the definition of the instantaneous energy (Eq C.2), yields

$$\frac{\partial E^p}{\partial w_{ij}(t)} = \frac{\partial E^p}{\partial c_k(t)}\, \delta_{jk}\, I_i \tag{C.11}$$

Now we must define the change in the energy of the network with respect to the actuator command at time t. The at any time (t+1), energy function can be described as being a function of the individual actuator commands at the same time.

$$E^p = E^p[c_1(t+1), c_2(t+1), \ldots, c_k(t+1)] \tag{C.12}$$

We know from Eqn 3.22, the definition of the i^{th} actuator command at time (t+1) is a function of the actuator commands one time step earlier, therefore the actuator command at time (t+1) can be defined as

$$c_k(t+1) = c_k(c_r(t))$$

Observe,

$$\frac{\partial E^p}{\partial c_k(t)} = \frac{\partial}{\partial c_k(t)} E^p \left[c_1(c_r(t), \ldots, c_k(c_r(t))\right] \tag{C.14}$$

Applying the chain rule to Eq (C.12) results in

$$\frac{\partial E^p}{\partial c_k(t)} = \sum_{r=1}^{K} \frac{\partial E^p}{\partial c_r(t+1)} \cdot \frac{\partial c_r(t+1)}{\partial c_k(t)} \tag{C.15}$$

The last partial derivative results in the feedback weight for the kith neuron. Eqn 3.50 can now be written as

$$\frac{\partial E^p}{\partial c_k(t)} = \sum_{r=1}^{K} \frac{\partial E^p}{\partial c_r(t+1)} T_{kr} \tag{C.16}$$

The change in weights for each iteration is now defined using Eqs (C.6, C.11, C.16) by

$$\Delta w_{ij}(t) = -\eta \left[\sum_{r=1}^{K} \frac{\partial E^p}{\partial c_r(t+1)} T_{jr}\right] \delta_{jk} I_i \tag{C.17}$$

If we look at the output when the neural network converges, the error function is described by

$$\varepsilon_k^p(t_e) = d_k - c_k(t_e) \tag{C.18}$$

The gradient of the energy function with respect to the input interconnection weights can be related to the instantaneous error by

$$\frac{\partial E^p}{\partial c_k(t_o)} = \frac{\partial}{\partial c_k(t_o)} \frac{1}{2} \sum_{k=1}^{K} (d_k - c_k(t_o))^2 \qquad \text{(C.19)}$$

Since the desired outputs remain fixed during convergence, the change in the desired output with respect to the weights is zero. The derivative is only defined when k=h, and therefore

$$\frac{\partial E^p}{\partial c_k(t_o)} = c_k(t_o) - d_k \qquad \text{(C.20)}$$

Now to the change in weights per iteration, we must begin after the network has converged and propagates backwards to t_o.

To summarize this procedure, the steps to train this type of recurrent network are as follows:

(1) Initialize weights, for this case the weights are initialized to the least square estimator ΔM_{jn} weight values to reduce training time.

(2) Present an input training exemplar and let network converge

(3) Using the number of iterations and Eqs C.19 and C.17, propagate backwards calculating the change in weights per iteration.

(4) Using Eqs C.5 and C.3, update weights.

(5) Present another training exemplar and proceed to Step 3

Bibliography

1. Aiyer, Sreeram V. B., *et al*, "A Theoretical Investigation into the Performance of the Hopfield Model," *IEEE Transactions on Neural Networks*, Vol. 1, No. 2, June 1990.

2. Angel J. R. P. *et al*, "Adaptive Optics for Array Telescopes using Neural-Network Techniques," *Nature*, Vol. 348, Nov 1990.

3. Chan L. W. and Fallside F., "An Adaptive Training Algorithm for Backpropagation Networks," *Computer Speech and Language*, Vol. 2, 1987.

4. Cochran G., "Phase Screen Generation," The Optical Sciences Company, *Report #663*, 1985.

5. Cohen Michael A. and Grossberg Stephen, "Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol SMC-13, No. 5, Sep/Oct 1983.

6. Ellerbroek, Brent L., "Wavefront and Tilt Estimators for the SOR-3 Experiment," *Report No. TR-684*, The Optical Sciences Company, December 1985.

7. Fisher, W. A. *et al*, "Application of the Lockheed Programmable Analog Neural Network Breadboard to the Real Time Adaptive Mirror Control Problem," *Proceedings of the International Society for Optical Engineering, Applications of Neural Networks*, Vol. 1294, 1990.

8. Fisher, W. A. *et al*, "A Programmable Analog Neural Network Processor," *IEEE Transactions on Neural Networks*, Vol. 2, No. 2, March 1991.

9. Fried D. L., "Statistics of a Geometric Representation of Wavefront Distortion," *Journal of the Optical Society of America*, Vol. 55, November 1965.

10. Gao Keqin *et al*, "A Neural Network Least-Square Estimator," *International Conference on Neural Networks, Proceedings*,

11. Gardner, Chester S *et al*, "Design and Performance Analysis of Adaptive Optical Telescopes Using Laser Guide Stars," *Proceedings of the IEEE*, Vol. 78, No. 11, November 1990.

12. Goodman, Joseph W., *Introduction to Fourier Optics*, McGraw-Hill Book Co., San Francisco, 1968.

13. Goodman, Joseph W.,*Statistical Optics*, Wiley-Interscience Publications, New York, 1985.

14. Hardy, John W., " Active Optics: A New Technology for the Control of Light," *Proceedings of the IEEE*, Vol. 66, No. 6, June 1978.

15. Hardy, John W., "Instrumental Limitations in Adaptive Optics for Astronomy," *Proceedings of the International Society of Optical Engineering, Active Telescope Systems*, Vol. 1114, March 1989.

16. Hardy, John W., "Real Time Atmospheric Compensation," *Journal of the Optical Society of America*, Vol. 67, No. 3, March 1977.

17. Hecht-Nielson, Robert, "Theory of the Backpropagation Neural Network," *International Conference on Neural Networks, Proceedings*, Vol. 1, September 1988.

18. Hopfield, John J., "Neurons with Graded Response have Collective Computational Properties like those of Two State Neurons," *Proceedings of the National Academy of Sciences*, Vol. 81, May 1984

19. Hudgin, Richard H., "Wave-front Reconstruction for Compensated Imaging," *Journal of the Optical Society of America*, Vol. 67, No. 3, March 1977.

20. Hudgin, Richard H., "Optimal Wavefront Estimation," *Journal of the Optical Society of America*, Vol. 67, No. 3, March 1977

21. Kosko, Bart, *Neural Networks and Fuzzy Systems*, Prentice Hall, New Jersey.

22. Lee, Bang W. and Sheu, Bing J., "Modified Hopfield Neural Networks for Retrieving the Optimal Solution," *IEEE Transactions on Neural Networks*, Vol 2, No. 1, Jan 1991.

23. Lippman, Richard P., "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine* , April 1987.

24. Melsa James L. and Cohn David L., *Detection and Estimation Theory*, McGraw-Hill Book Company, New York.

25. Noll, Robert J., "Zernicke Polynomials and Atmospheric Turbulence," *Journal of the Optical Society of America*, Vol. 66, January 1976.

26. Paxman R. G. and Fienup J. R., "Optical Misalignment Sensing and Image Reconstruction using Phase Diversity," *Journal of the Optical Society of America*, Vol. 5, No. 6, June 1988.

27. Rogers, Steven K. *et al*, *Introduction to Biological and Artificial Neural Networks*, Air Force Institute of Technology, October 1990.

28. Roggemann, Michael C., "Optical Performance of Fully and Partially Compensated Adaptive optics Systems Using Least-Squares and Minimum Variance Phase Reconstructors," *Preprint: Computers and Electrical Engineering.*

29. Sandler D. G. *et al*, "Use of a Neural Network to Control an Adaptive Optics System for an Astronomical Telescope," *Nature*, Vol. 351, May 1991.

30. Sheu Bing J. *et al*, "Hardware Annealing for Fast Retrieval of Optimal Solutions in Hopfield Neural Networks," *IEEE* ,

31. Smithson, Robert C., "A Segmented Mirror for Solar Observation," *Proceedings of the International Society for Optical Engineering, Electromechanical System Interaction with Optical Design*, Vol. 779, March 1987.

32. Takeda Mitsuo and Goodman Joseph W., "Neural Networks for Computation: Number representations and Programming Complexity," *Applied Optics*, Vol. 25, No.18, Sep 1986.

33. Tank, David W. and Hopfield John J., "Simple "Neural" Optimization Networks: An A/D Converter, Signal Decision Circuit, and a Linear Programming Circuit," *IEEE Transactions on Circuits and Systems*, Vol. CAS-33, No. 5, May 1986.

34. Tatarski, V. I., *Wave Propagation in Turbulent Media*, Dover Publishers, New York, 1967.

35. Wallner, Edward P., "Minimizing Atmospheric Dispersion Effects in Compensated Imaging," *Journal of the Optical Society of America*, Vol. 67, No. 3, March 1977.

36. Wang, J. Y., "Phase Compensated Optical Beam Propagation through Atmospheric Turbulence," *Applied Optics*, Vol. 17, August 1978.

37. Wasserman, Philip D., *Neural Computing: Theory and Practice*, Van Nostrand Reinhold, New York

38. Welsh Byron M. and Gardner Chester S., "Performance Analysis of Adaptive-Optics Systems Using Laser Guide Stars and Slope Sensors," *Journal of the Optical Society of America*, Vol. 6, No. 12, December 1989

39. Williams, Ronald J. and Zipser, David, "A Learning Algorithm for Continually Running Fully Recurrent Neural Networks," *Neural Computation*, Vol. 1, 1989.

40. Zahirniak, Daniel R., "*Characterization of Radar Signals Using Neural Networks*," MS Thesis AFIT/GE/ENG/90D-69. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, December 1990.

41. Zhao Xiaofeng, and Mendel Jerry M., "An Artificial Neural Minimum-Variance Estimator," *Proceedings: International Conference on Neural Networks*,